

State of Arizona



Arizona Broadband Assessment Project (AZ BAP)

Methodology White Paper

Submission 6 - October 1, 2011

State of Arizona

Arizona Strategic Enterprise Technology Office (ASET)

Arizona Broadband Assessment Project Methodology White Paper

Submission 6 - October 1, 2012

Arizona Broadband Assessment Project Overview

The purpose of the Arizona Broadband Assessment Project (AZ BAP) is to identify both the availability and speed of broadband services, and the location of broadband infrastructure throughout Arizona. This project is provided through the American Recovery and Reinvestment Act of 2009 (ARRA) and the Broadband Data Improvement Act (BDIA), and in conjunction with the National Telecommunications and Information Administration (NTIA).

Submission 6 for the broadband availability data set was duly submitted to NTIA prior to October 1, 2012. Fall 2012 was the sixth of ten semi-annual submissions by the State of Arizona and attempts to capture and reflect broadband availability and conditions in the field as of June 30, 2012. See Arizona's Changes & Corrections document for a list of included Broadband Providers and relevant notes.

The Methodology White Paper for this submission cycle focuses on the Data Intake and Validation Application (DIVA) developed by the Arizona team and utilized in the transformation and processing of Broadband Provider data. It was trialed in earlier submittals, but has reached a level of maturity and utility warranting its featured role here.

A brief overview of its development and role as well as a summary of its capabilities and impact, followed by a draft of the DIVA User Manual for a deeper view of its functionality and user interface.

Data Intake and Validation Application (DIVA) Overview

The Data Intake and Validation Application (DIVA) was developed by TerraSystems Southwest (TSSW) as a subcontractor to Data Site Consortium (DSC) under contract with Arizona Strategic Enterprise Technology Office (ASET) in support of the Arizona Broadband Assessment Project (AZ BAP). DIVA is a Windows desktop application designed to transform raw Broadband Provider data about the location, technology and speed of broadband services into a form that can be cleanly linked to GIS layers and imported into the NTIA standard national broadband mapping program geodatabase.

A key goal of the DIVA design was to reduce data processing time while increasing data integrity. A secondary goal was to create a freely distributable software tool that Providers and other State broadband organizations could apply to their data intake and validation tasks. Alternative approaches, such as integrating with ArcGIS or data translation software like FME

were not pursued as that would mean users would have to purchase those products at a significant expense.

DIVA does not perform any spatial validation or processing. DIVA was scoped as “pre-GIS” software, designed only to speed and improve the processing of Provider data to a point where it could be more cleanly geocoded or linked to NTIA GIS layers. Based on this design criteria, DIVA is not very useful for wireless service shapefile deliveries where the feature counts (data Records) are a couple of hundred, or less, and are in more or less proper SBDD format. It really shines in processing address, census block and road segment submittals of tens or hundreds of thousands of records, and where Providers have not followed the SBDD coding scheme.

Data Intake and Validation Application (DIVA) Capabilities

DIVA offers a rich user interface for exploring and processing Provider broadband data into a form that can more easily linked to NTIA-required GIS feature classes. Some of its key features are listed below.

- **Configurability** – DIVA offers many opportunities for configuration. New Provider identification information can be imported and applied to every Provider submittal. New releases of the SBDD geodatabase are read and up-to-date Rules are automatically created and applied. Processes and Rules in DIVA are very general and may be user-configured to achieve various results.
- **Consistency** – a very structured approach to data processing is embedded in the design of DIVA. This begins with the clear definition of data elements and their relationship to one another in an Object Data Model. Consistency is also inherent in the clear definition of Processes and Rules that can be applied to the data and in the way that Processes and Rules are used to transform and validate the output data. In the rush of meeting data-delivery deadlines It is easy to forget or misapply data processing steps. By automating much of the required processing, DIVA increase the amount of time that a user has to actually review and check data, and makes it easier for the user to achieve consistent results in the exported data sets.
- **Re-usability** – Users can define a set of Processes within a particular Reporting Period as a Template and then apply the Template to new Input Files. Rules are uniformly applied to Providers for each Reporting Period. This includes user-defined rules: once defined and applied, they will automatically be applied to subsequent Submittals.
- **Processing Documentation** – metadata (e.g. notes) regarding Providers, Submittals, Input Files and other elements may be added at any time using the “Edit Metadata” button on the Status Bar. Notes can be viewed or exported at any time for cutting and pasting into NTIA documentation. These notes, plus the actual Input File(s) associated with a Submittal, the assigned Processes, Rule violations and final output, constitute DIVA’s Metadata system. A good example of metadata stored in a Process is the translation table from Provider actual speed values to NTIA speed tier codes: the value mapping is preserved and can be reviewed in DIVA by opening up the applicable Submittal and generating a detailed Input File report or by right-clicking on the Translation Process in the Processes tab.

Data Intake and Validation Application (DIVA) Impact

The efficiencies resulting from the application of DIVA to Provider data are substantial. In the first submittal period in Spring of 2010 a number of larger providers would consume 24-40 hours of processing time to evaluate, transform, quality check and export to SBDD database format. Processing a similar set of data in the latest submittal period (Spring 2012) using DIVA is a 2-4 hour process.

A substantial portion of this improvement is the result of (a) knowing the data and what to expect from a given Provider and (b) improved manual processing, especially on the GIS side. However, another substantial portion of the improvement has come from the integrated data evaluation, checking, transformation, validation and export capabilities of DIVA itself. We estimate that DIVA can reduce processing time in half for large address or census block submittals from the Providers. This efficiency is gained from having all the evaluation, transformation and validation tools available in a single interface instead of applying a variety of application software packages in varying order to each Provider file in each Reporting Period.

Arizona Broadband Assessment Project (AZ BAP)

The Data Intake and Validation Application (DIVA)

User Manual for Version 1.0

As of 06/25/2012

The Data Intake and Validation Application (DIVA)

Table of Contents

[Foreword](#)

[DIVA Overview](#)

[History](#)

[Design Goals](#)

[The DIVA “Paradigm”](#)

[Generalized Processing Workflow](#)

[Key Features](#)

[What DIVA is Not](#)

[Installation](#)

[Functionality Overview by GUI Elements](#)

[Status Bar Overview](#)

[Submittals Tab Overview](#)

[Processes Tab Overview](#)

[Rules Tab Overview](#)

[Output Tab Overview](#)

[Utilities Tab Overview](#)

[Reports Tab Overview](#)

[About Tab Overview](#)

[Functionality Details](#)

[Status Bar Functions](#)

[Submittals Tab Functions](#)

[Processes Tab Functions](#)

[Rules Tab Functions](#)

[Output Tab Functions](#)

[Utilities Tab Functions](#)

[Reports Tab Functions](#)

[Sample Workflow](#)

[Notes on Bugs and Desired Functionality](#)

[Glossary](#)

[Appendix A: User Needs Analysis](#)

[Appendix B: Guidelines for Pre-Processing Provider Data](#)

[Appendix C: Object Data Model](#)

[Appendix D: Changes DIVA Makes in SQL Server](#)

Foreword

This User Manual includes three chapters to introduce users to Diva. The “[Functionality Overview by GUI Elements](#)” chapter shows keyed illustrations of the user-accessible controls on each major Graphical User Interface (GUI) element in DIVA along with a brief explanation of the purpose of those controls. More detailed descriptions of the functionality behind many of those controls can be found by clicking on the hyperlinks which take the user to the appropriate section in the “[Functionality Details](#)” chapter. Finally, an example workflow using DIVA is presented in the “Sample Workflow” chapter.

It is important before examining the use of DIVA in detail, however, that the user understands the overall approach that DIVA takes to processing data. The discussion in “[The DIVA ‘Paradigm’](#)” section of the “[DIVA Overview](#)” chapter is the key to understanding this approach. Also important to user understanding are the summaries in the “[Design Goals](#)”, “[Processing Workflow](#)” and “[Key Features](#)” sections of that chapter.

As part of the DIVA development process, a very specific, structured vocabulary was developed and is enforced throughout this document. The first time a particular term is used, it is hyperlinked to the [Glossary](#) where a complete definition of the term may be found. Users will greatly speed their understanding of DIVA if they become familiar with this terminology.

DIVA is made available to the State Broadband community at no charge, but it is not supported by TerraSystems or any other public or private organization. The developers are happy to hear comments, bug reports and suggestions for improvement via email at hlward@terrasw.com , but we cannot guarantee a response to each inquiry. The source code is available to the State broadband community under certain circumstances. Please ask for details at the email address above.

DIVA Overview

History:

The Data Intake and Validation Application (DIVA) was developed by TerraSystems Southwest as a subcontractor to Data Site Consortium under contract with the Arizona Broadband Mapping program.

DIVA is a Windows desktop application designed to process raw Broadband [Provider](#) data about the location, technology and speed of broadband services into a form that can then be cleanly linked to GIS layers and imported into the NTIA standard national broadband mapping program geodatabase.

Design Goals:

A key goal of the DIVA design was to reduce data processing time while increasing data integrity. A secondary goal was to create a freely distributable software tool that Providers and other State broadband organizations could apply to their data intake and validation tasks. Alternative approaches, such as integrating with ArcGIS or data translation software like FME, would mean users would have to purchase those products at a significant expense.

By design, DIVA only processes provider data to the point where it can be geocoded or joined to spatial database layers. It does not perform any spatial validation functions. See [Appendix A](#) for a summary of the User Needs document on which DIVA was designed.

The DIVA “Paradigm”:

Using DIVA effectively requires an understanding of how DIVA approaches the processing of Provider data for the NTIA and, as noted in the Foreword, this view is expressed with specific vocabulary. The DIVA “Paradigm” will be outlined in these terms. Links to the glossary can be used for further reference. Figure 1 below illustrates the main DIVA entities (the “object model”) and their relationships in a simplified manner. An illustration of the complete object model can be found in [Appendix C](#).

A [Submittal](#) is the data that will be reported to the NTIA for a particular [Provider](#) for a particular [Period-Type \(Reporting Period / Submittal Type\)](#). **The Submittal is the key organizing element in DIVA.** The collection of all Submittals is what will ultimately be reported to the NTIA. One of the important consequences of this

definition is that, for one Provider, there may be several Submittals in a Reporting Period, e.g. Census Block and Street Segment submittals.

It is possible that a Provider will provide more than one data set for a particular Submittal Type during a Reporting Period. For example, supplementary information may be sent by a Provider after their first provision of data, or the first provided data set might be entirely replaced by another. DIVA identifies each of these provided data sets as a separate [Input File](#), and allows a Submittal to have more than one Input File in a Reporting Period. Each Input File is processed independently, although the resulting exported data may be combined later, outside of DIVA, depending on the reasons for the multiple data set submission.

An Input File in DIVA is a collection of [Input Fields](#) each of which (usually) has a value. This Input Field collection is repeated numerous times, each repetition being viewed as a data [Record](#) or [Row](#). The order of the Input Fields in this collection does not change from Record to Record. Each Input Field may be viewed as a data [Column](#) when it is considered across all data Records (Rows).

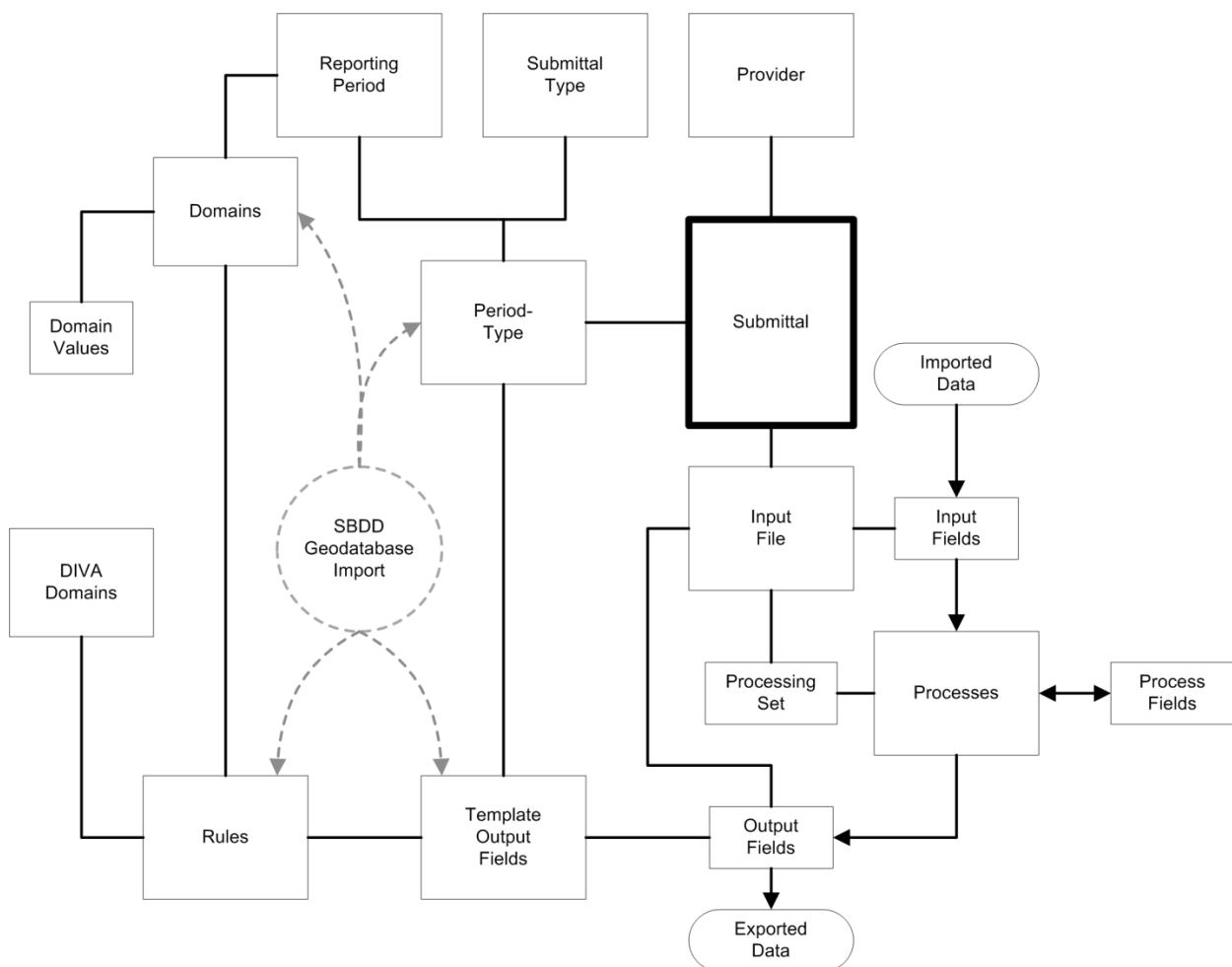


Figure 1. The DIVA “Paradigm” (Object Model) in simplified form.

Lines denote relationships and arrows denote data flow.

Each [Reporting Period](#), the NTIA requires that certain data be submitted for each participating Broadband [Provider](#) in the State. The data is grouped into particular [Submittal Types](#) (e.g., service by Census Blocks) and these types and the required information for each type, the [Output Fields](#), are defined by the NTIA prior to each Reporting Period in the form of an [SBDD ArcGIS File Geodatabase](#) (SBDD geodatabase, for short). A Submittal Type corresponds with a feature class in the geodatabase and the Output Fields are the attribute fields of each feature class. Since the same Output Fields will be used with each Submittal Type for a Reporting Period (i.e., for each Period-Type) for every Provider, they actually constitute a set of [Template Output Fields](#) that are used to create each particular Submittal's Output Fields that are eventually exported from DIVA. Thus, importing the SBDD geodatabase results in the creation of a set of Period-Types and a set of Template Output Fields for each Period-Type. These Template Output Fields will ultimately be used to create a set of Output Fields for each Input File.

In addition, the SBDD geodatabase contains a set of [Domains](#) that define the permissible values for a particular Template Output Field; the SBDD geodatabase specifies the Domain names and values as well as the Template Output Field to which it applies. Since the names and / or contents of the Domains may vary for each Reporting Period, a new set of Domains and values is created during SBDD geodatabase import and is used only for that Reporting Period.

In some cases, the SBDD geodatabase will specify a [Default Value](#) for an Output Field. Since this applies to all Output Fields for that Period-Type, the value is stored as part of the Template Output Field definition during SBDD geodatabase import.

While the feature classes in the SBDD geodatabase have been fairly consistent over all Reporting Periods, their defined attributes (Template Output Fields) have not been, nor have the number of Domains, their values, or assignment to Template Output Fields remained unchanged. DIVA handles this variability by creating a match between a Reporting Period and a Submittal Type called a Period-Type. Thus the "Spring 2011 Census Block" Period-Type may have a different structure (Template Output Fields, Domains, Domain assignments, etc.) than the "Spring 2012 Census Block" Period-Type. So when we speak of a "Census Block" Submittal Type, it is usually understood that this is for a particular Reporting Period and that we are actually referring to the matching Period-Type.

The "SBDD Geodatabase Import" operation, which must be done once for each Reporting Period, creates all of the Reporting Period's Period-Types (one for each feature class in the geodatabase, each feature class corresponding to a Submittal Type), creates each Period-Type's Template Output Fields (one for each attribute in a feature class, plus some additional DIVA-specific fields), and creates a set of "default" [Rules](#) that are attached to the Template Output Fields. It also assigns Default Values, if defined, to the Template Output Fields.

DIVA uses pre-defined, general data transformation operations, called [Processes](#), to convert Input Field values to Output Field values. The collection of all Processes for a particular Input File is held in a [Processing Set](#), and this set may be copied to other Input Files or used to create a [Template Processing Set](#). As can be seen in Figure 1, data moves from an Input Field to an Output Field only by using a Process. In addition, depending on how Processes are “wired” together, intermediate [Process Fields](#) may be created. This is done according to the operational requirement that Processes cannot connect directly to each other (connections can only be made to Fields) and Fields cannot connect directly to each other (connections can only be made to Processes).

One of the primary reasons for creating DIVA was to automate, to a certain degree, the process of checking data. DIVA handles both [data verification](#) and [data validation](#) (the term “validation” is generally used to mean both of these, but we use these terms in a very particular sense; see Glossary). The DIVA Rule is our way of implementing this checking. For example, if an Output Field is required to have data according to the SBDD Geodatabase, then a “MustHaveData” rule is assigned to the corresponding created Template Output Field. Since all Output Fields for an Input File (Submittal) of a particular Period-Type are derived from the same set of Template Output Fields, these Rules thus apply to all such Output Fields. Rule violations can be checked after Processes are applied, and then the Processes can be adjusted to correct the data transformations to reduce the number of Rule violations.

An important item to note regarding DIVA data processing is that, unlike Input Field values, **Output Field values are never explicitly stored (saved) in an internal data table.** Output Field values are only generated during Process operations; the Processes tab must be “visited” in order to generate any Output Field values. This has important consequences in the order in which various Tabs should be visited, as well as to what information will be shown in various display windows. For instance, selecting a Submittal and then immediately opening the [“Output Data Review”](#) window will result in no output data being displayed; it has not yet been generated. You must click on the Processes tab to generate the data (or, if you are creating Processes on that Tab, then you must click the “Apply” button to have any changes take effect). If the Output Data Review window is open at this time, you will see the generated data “fill” the display.

Because Output Fields will not have values until the Processes tab is visited, the Output tab will not have any meaningful results, in terms of checking for Rule violations, until the Processes tab is visited.

The summary and detailed Input File reports will not have meaningful data until both the Processes tab is visited and the “Refresh Violation Counts” button is clicked on the Output Tab.

Generalized Processing Workflow:

The general DIVA processing workflow for a particular set of data (an Input File) for a Submittal is as follows:

1. Submittals tab: Create the Submittal or add an Input File to an existing Submittal.
2. Submittals tab: Load data from a Provider's data set (source file on disk) into the Input File (stored in an internal DIVA data table).
3. Processes tab: Create and apply transformative Processes linking Output Fields to a data source. A data source is typically an Input Field, but some Processes can serve as a data source themselves.
4. Rules tab: Assign Rules to check values in the Output Fields.
5. Output tab: Examine data errors.
6. Iterate over steps 3 – 5 to fix errors in the Output Fields by modifying assigned Processes and / or adjusting Rules as necessary.
7. Output Data Review window: Examine and export the transformed data.
8. Reports tab: Generate reports.

Key Features:

DIVA offers a rich user interface for exploring and processing Provider broadband data into a form that can more easily linked to NTIA-required GIS feature classes.

- **Configurability** – DIVA offers many opportunities for configuration. New Provider identification information can be imported and applied to every Provider submittal. New releases of the SBDD geodatabase are read and up-to-date Rules are automatically created and applied. Processes and Rules in DIVA are very general and may be user-configured to achieve various results.
- **Consistency** – The “structured approach” to data processing is embedded in the design of DIVA: in the Object Data Model that defines DIVA entities, in the definition of Processes, in the definition of Rules, and in the ways that Processes and Rules are used to generate and check output data. It is easy to lose data integrity in the rush of meeting data-delivery deadlines. By automating much of this processing, DIVA tries to increase the amount of time that a user has to actually review and check data, and made it easier for the user to achieve consistent, known results in the exported data sets.
- **Re-usability** – Users can define a set of Processes within a particular Reporting Period as a Template and then apply the Template to new Input Files. Rules are uniformly applied to Providers for each Reporting Period. This includes user-defined rules: once defined and applied, they will automatically be applied to subsequent Submittals.
- **Documentation** – Notes regarding Providers, Submittals, Input Files and other elements may be added at any time using the “Edit Metadata” button on the Status Bar. Notes can be viewed or exported at any time for cutting and pasting into NTIA documentation. These notes, plus the actual Input File(s) associated with a Submittal, the assigned Processes, Rule violations and final output, constitute DIVA’s Metadata system. A good example of metadata stored in a Process is the translation table from Provider actual speed values to NTIA speed tier codes: the value mapping is preserved and can be reviewed in DIVA by opening up the applicable Submittal and generating a detailed Input File report (or by right-clicking on the Translation Process in the Processes tab).

What DIVA is Not:

DIVA does not perform any spatial validation or processing. DIVA was scoped as “pre-GIS” software, designed only to speed and improve the processing of Provider data to a point where it could be more cleanly geocoded or linked to NTIA GIS layers. Based on this design criteria, DIVA is not very useful for wireless service shapefile deliveries where the feature counts (data Records) are a couple of hundred, or less, and are in more or less proper SBDD format. It really shines in processing address, census block and road segment submittals of tens or hundreds of thousands of records, and where Providers have not followed the SBDD coding scheme.

Installation

DIVA comes with a basic installer which installs the application program, builds the basic SQL Server schema and populates various tables with DIVA-standard values (e.g. Reporting Period names). The DIVA installation requires and assumes you have already installed the latest free version of Microsoft's SQL Server Express 2008 R2.

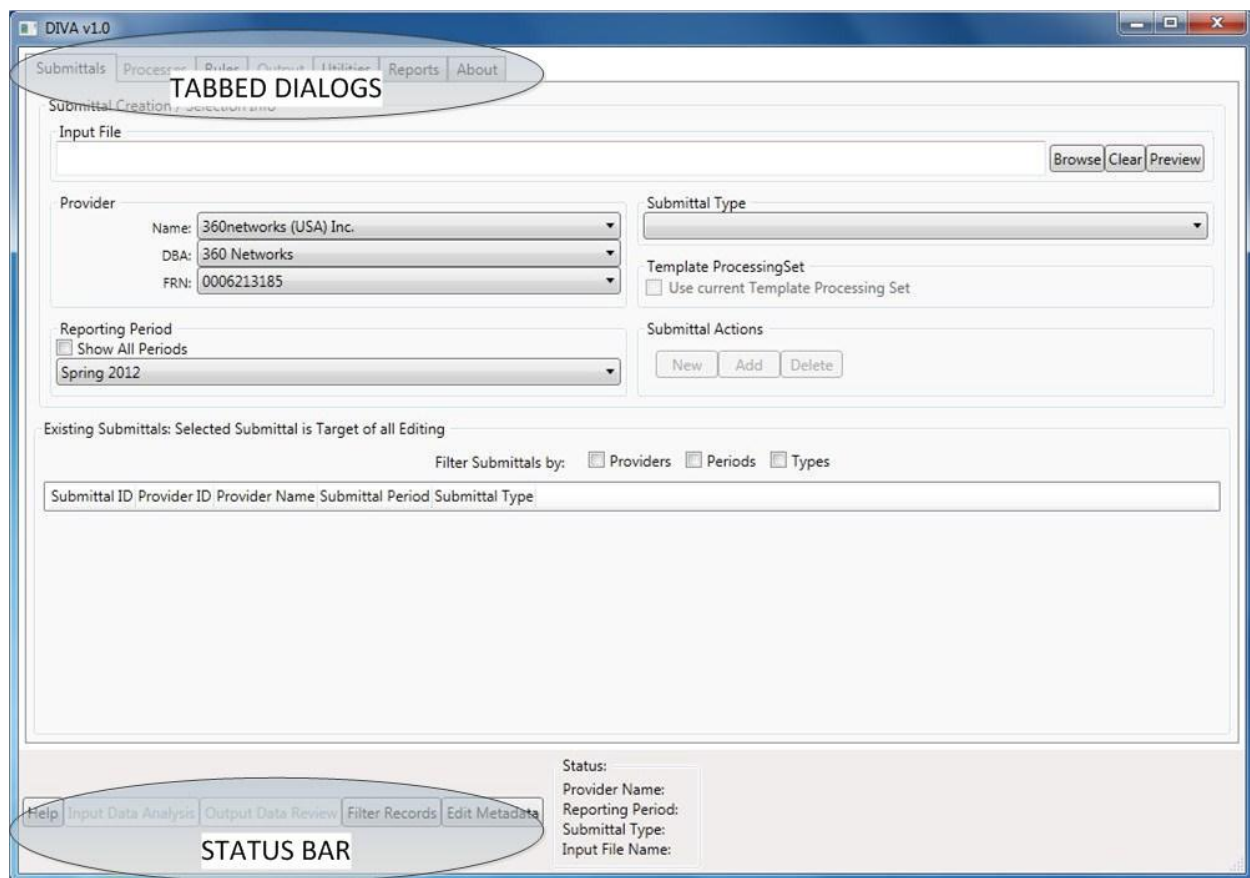
Basic installation steps are:

1. Create an install folder on your system
2. Copy the "DIVA.exe.config" file from the previous prototype folder to the new install folder.
3. In SQL Server, delete and then re-create DIVA database.
4. In SQL Server, run "DivaModel.edmx.sql"
5. In SQL Server, run, "BaseData.sql "

Provider data will now be loaded into period-specific data bases. THESE WILL HAVE TO BE DELETED IN THE FUTURE as part of the re-install. Their names look like "Diva_PData_X##" where "X" is "F" or "S" and "XX" is the two digit year. For a description of the tables that DIVA creates in SQL Server please see [Appendix D](#).

Functionality Overview by Graphic User Interface (GUI) Elements

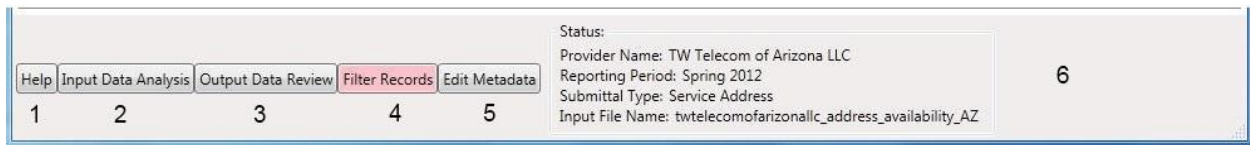
The main DIVA interface is a tabbed dialog box with the tabs organized from left to right roughly corresponding to the order in which a user would proceed while processing Provider data. The image below shows the initial DIVA view, the Submittals Tab. In addition to the left-to-right ordering of tabs, DIVA also guides users by activating tabs and various controls only when pre-requisite steps have been completed. This guidance is only approximate. DIVA was not designed as a data processing “wizard” and therefore users are allowed to choose many pathways through the DIVA application. In addition to tabs, DIVA has a “Status Bar”, composed of five buttons and a status display area found at the bottom of the DIVA GUI regardless of which tab is currently active.



Following are brief explanations of the tabs and Status Bar controls on the DIVA GUI. Hyperlinks reference more detailed descriptions of how and when to apply a particular control.

Status Bar Overview

For the most part, buttons on the Status Bar can be used at any time, which is why they are visible no matter which tab is active. The “Status” area on the right side of the bar displays the currently [active Provider](#), Period-Type (Reporting Period and Submittal Type) and Input File Name. Following is a table that briefly describes the function of each of the Status Bar elements; if an element is a button, then the description is what happens when the button is clicked. Follow hyperlinks for more detailed information.



Key	Control	Description
1	Help button	Brings up an HTML version of selected portions of this documentation.
2	Input Data Analysis button	Opens the “Input Data Analysis” window which displays the data that was loaded into DIVA from the active Input File for the active Submittal.
3	Output Data Review button	<p>Opens the “Output Data Review” window and displays the data resulting from applying the Processes (as configured on the “Processes” tab) to the Input Fields and the errors resulting from applying Rules on the Output tab.</p> <p><i>Note: after selecting a Submittal, users must click on the “Processes Tab” to apply the Processes for the selected Submittal. This will generate the Output Field data and the “Output Data Review” window will be populated with those values.</i></p> <p><i>Note: after applying Processes, users must click on the “Refresh Violation Counts” on the Output tab. Errors will then be viewable in this window.</i></p>
4	Filter Records button	<p>Opens the “Filter Records Window” where the user can specify a number of records or an SQL query expression (or both) to restrict the number of records that are loaded for current processing.</p> <p><i>Note that, in the example above, the button is colored. This indicates that a filter is in effect (i.e., not all Input File records are currently being processed).</i></p>

Key	Control	Description
5	Edit Metadata button	Opens the “Edit Metadata Window” where users can record notes about Providers, Reporting Periods, Submittal Types and Submittals. Metadata entries are included, as appropriate, in the various Reports.
6	Status area	The “Status” area is always visible and reports on information about the active Submittal.

Submittals Tab Overview

The Submittals tab contains all the controls a user needs to manage the intake of provider data. In the image below are numeric keys to the discussion that follows. More detailed information on how and when to use these controls can be found in the “Functional Details” and “Sample Workflow” sections of this documentation.

The upper portion of the tab, the “Submittal Creation / Selection Info” panel, is used to set the [active](#) Provider, Reporting Period and Submittal Type. If this combination has been used previously to define a Submittal, then that Submittal will automatically be selected in the lower panel and the selected data set (the Input File) can be added to the Submittal. If this combination has not been used previously, then a new Submittal (with the selected Input File data set) can be created.

The lower portion of the Tab, the “Existing Submittals” panel, displays all existing Submittals and can also be used to navigate among them. The “Filter Submittals” check boxes allow the user to restrict the display of existing Submittals to only those matching the active Provider, Reporting Period and / or Submittal Type.

DIVA v5.2

Submittals | Processes | Rules | Output | Utilities | Reports | About

Submittal Creation / Selection Info

Input File: H:\Broadband\N03_Data_Processing\aa2012_01_Spring\TW_Telecom\TW_NearResults.txt (1) [Browse] [Clear] [Preview] (2, 3, 4)

Provider (5): Name: TW Telecom of Arizona LLC, DBA: TW Telecom of Arizona LLC, FRN: 0004352274

Submittal Type: Service Address (6)

Template Processing Set: [] Use current Template Processing Set (7)

Reporting Period: [x] Show All Periods, Spring 2012 (8)

Submittal Actions: [New] [Add] [Delete] (9)

Existing Submittals: Selected Submittal is Target of all Editing

Filter Submittals by: [] Providers [] Periods [] Types

Submittal ID	Provider ID	Provider Name	Submittal Period	Submittal Type
2	TW	TW Telecom of Arizona LLC	Spring 2012	Service Address
6	ME	Mediacom South East LLC	Spring 2012	Service Address

Selected Submittal's Input File Information

twtelecomofarizonallc_address_availability_AZ (11) [Delete] (12) [Load] (13) Processing Set: TW_S12_AD_2 (14) [Copy] [Paste]

Status:
Provider Name: TW Telecom of Arizona LLC
Reporting Period: Spring 2012
Submittal Type: Service Address
Input File Name: twtelecomofarizonallc_address_availability_AZ

[Help] [Input Data Analysis] [Output Data Review] [Filter Records] [Edit Metadata]

Key	Control	Description
1	Input File text box	Displays the path and filename of the Provider Data Set which a user has selected using the Browse button. <i>Note: more than one Input File may be associated with a submittal. These files are processed separately and not combined by DIVA in any way.</i>
2	Browse button	This button opens a file dialog box which is used to select a Provider data set (a DIVA Input File).
3	Clear button	Clears the Input File selection from the Input File text box.
4	Preview button	Opens the selected Input File in an “Input Preview” dialog box, allowing the user to evaluate the content and format of the first 100 records of the Input File.
5	Provider area	The three drop-down lists in this area allow the user to set the active Provider using either the Provider Name, DBA Name or FRN (FCC Registration Number). Selection in one drop-down list automatically synchronizes the other two. Users may also click on an existing Submittal in the Existing Submittals list to set the active Provider (i.e., display the Provider’s identification information). The Provider drop-down lists are populated from an internal DIVA table. This table can be updated from the Utility tab (Update Provider Info).
6	Reporting Period drop-down list	This drop-down list is used to set the active Reporting Period. Only the current Reporting Period is available for selection unless the “Show All Periods” box is checked. Each Reporting Period has an internal DIVA-defined Start and End date. The Reporting Period that has a date range containing the current system date is considered current. The Reporting Period pull-down list can be modified from the Utility tab (Manage Submittal Reporting Periods).
7	Submittal Type drop-down list	This drop-down list is used to set the active Submittal Type. Each of the eight NTIA approved Submittal Types <u>may</u> be listed. The ones that <u>will</u> be shown are the ones for which a matching feature class was found during SBDD XML Import for this Reporting Period. <i>Note: This is used in combination with the Reporting Period drop-down list to set the active Period-Type.</i>

Key	Control	Description
8	Template Processing Set checkbox	<p>Checking this box, prior to clicking on the Processes tab for any new Submittal, will automatically generate Processes for that Submittal from the current Template Processing Set (if it exists) when the Processes tab is clicked.</p> <p>Template Processing Sets are created on the Processes Tab. Any existing Processes for that Submittal will be copied to the Template Processing Set overwriting any previously saved Processing Set for the current Period Type.</p> <p><i>Note: there can be one and only one Template Processing Set for each Submittal Type for a given Reporting Period, that is, for any one Period-Type.</i></p>
9	Submittal Actions buttons	<p>The “New”, “Add” and “Delete” buttons in this section of the Submittals tab all pertain to working with Submittals.</p> <p>The “New” button will only be available if no previous Submittal has the same assigned Provider, Reporting Period and Submittal Type. Clicking the button will create the Submittal with these values assigned and the currently selected data set as the Input File. This new Submittal will be added to the “Existing Submittals” list and it will be active (selected).</p> <p>The “Add” button will only be available if the active Provider, Reporting Period and Submittal Type have been used previously (i.e., there exists a Submittal with these values). In this case, the selected Input File will be added to that Submittal; it will be added to the Existing Input Files drop-down list and be made the active Input File.</p> <p>The “Delete” button will completely delete an existing Submittal, including references to its Input File(s), their Input and Output Fields, and their Processing Sets (any defined Processes and Process Fields). The internally stored data of the Input File(s) are also deleted.</p> <p><i>Note: each new data set added to an existing Submittal is its own Input File. In other words, multiple Input Files are not appended and treated as a single Input File.</i></p>

Key	Control	Description
10	Existing Submittals list	<p>This is a list of <u>all</u> Submittals that users have previously defined using DIVA. Clicking on a Submittal in this list makes it the active Submittal and synchronizes the information in the upper panel of the Submittals tab (the “Provider”, “Reporting Period” and “Submittal Type”) with the active Submittal.</p> <p>The checkboxes along the top filter the list of existing Submittals found in the Existing Submittals table area. Users may filter the list to specific Providers, Reporting Periods and / or Submittal Types.</p> <p><i>Note: When a Submittal is created, the Input File’s data is not automatically loaded into DIVA’s internal storage. Users may have referenced several Input Files for a Submittal without loading data. Data loading is a separate step from Submittal creation and is required for any further processing in DIVA.</i></p>
11	Input File drop-down list	<p>Displays the active Input File name for the selected Submittal. If there are multiple Input Files assigned to a Submittal, the active Input File may be selected using this drop-down list.</p> <p><i>Note: When a Submittal that has multiple Input Files is selected, the first Input File will be active (selected). The user must change this as needed. DIVA does <u>not</u> remember (and restore) the last Input File that was used for this Submittal.</i></p>
12	Delete button	<p>Deletes the active (selected) Input File. This deletes all associated information in the DIVA database, including Processes and Metadata, for that Input File <u>only</u>; the selected Submittal is not deleted. If there is only one Input File for a Submittal, then it is not possible to delete it and the button is inaccessible; use the Submittal Delete button instead.</p> <p><i>Note: A Submittal must have at least one Input File; therefore, to remove the last Input File the user must delete the entire Submittal.</i></p>
13	Load button	<p>Opens a “Load Data” dialog box (this is very similar to the “Input Preview” dialog box accessed with the Preview button) and allows the user to evaluate the content and format of the first 100 records of the Input File. The settings here are more critical because, when the Load button in this dialog box is clicked, these settings will determine how the data file will be loaded into DIVA’s internal database.</p> <p><i>Note: It is important to understand that you can Create a new Submittal, but nothing further can be done with it until you load the data from the Input File.</i></p>

Key	Control	Description
14	Processing Set area	<p>This area shows the (internal) name of the Processing Set attached to the active (selected) Input File, and contains buttons for copying and pasting a Processing Set to another Submittal's Input File.</p> <p>The "Copy" button will be accessible only if this Input File has had its data loaded <u>and</u> its Processing Set contains at least one Process (the user must have created and saved at least one Process on the Processes Tab). Clicking the Copy button will "tag" this Processing Set for copying.</p> <p>The "Paste" button will be accessible only if this Input File has had its data loaded, another Submittal has had its Processing Set "tagged" for copying, <u>and</u> this Input File's Processing Set is <u>empty</u> (no Processes ever assigned to it). The user must remove all Processes from an Input File (by deleting them on the Processes tab and then saving the results) if another Processing Set is to be "pasted". Clicking the Paste button will automatically transfer DIVA "focus" to the Processes tab so that the Paste operation can be completed.</p>

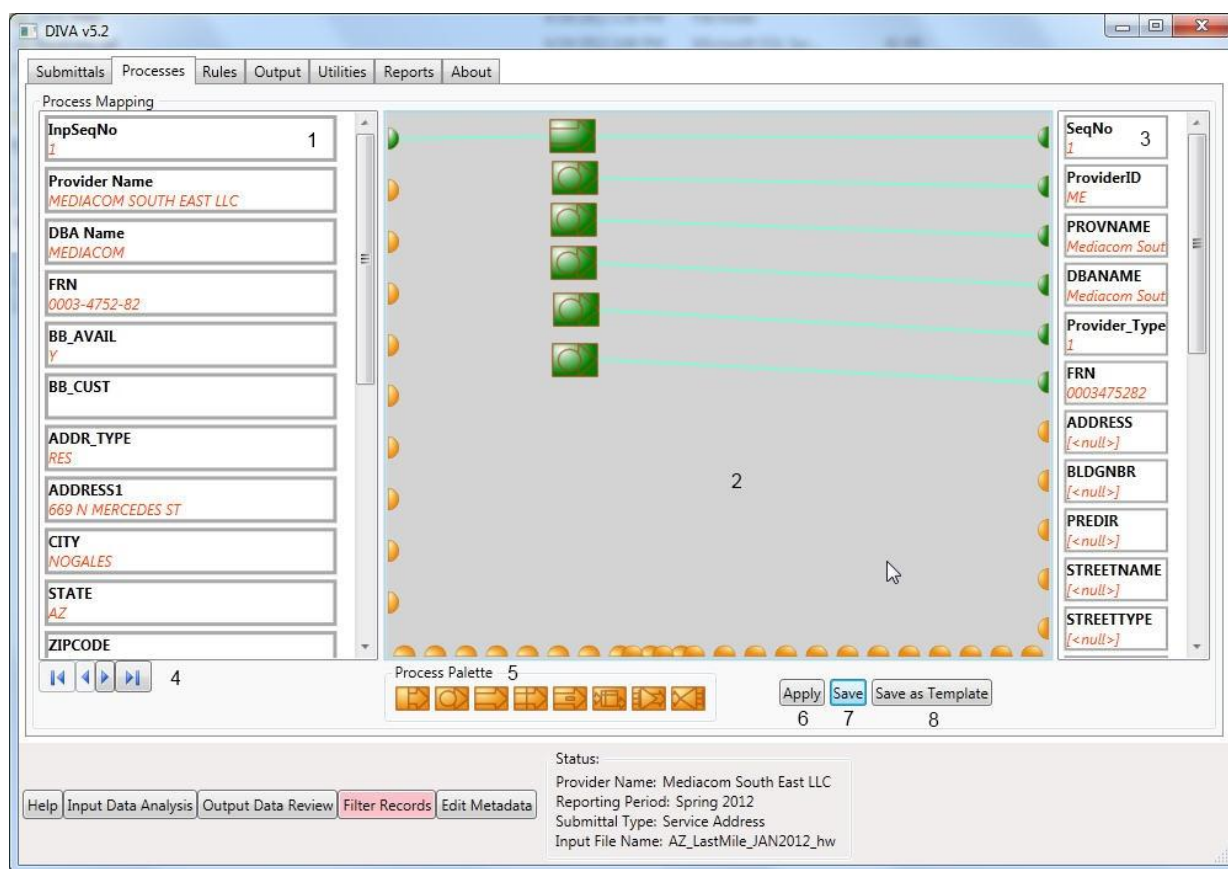
It is important to note that once a Provider's data set (computer file contents) has been loaded into internal DIVA storage (as a data table in a Reporting Period-specific database), that data set's computer file may be renamed, moved or deleted with no effect on DIVA. Prior to loading, changing the filename, the file's location or deleting the file will cause the DIVA data loading operation to fail. Doing any of these actions will force the user to browse to a new data set and add it as an Input File for that Submittal or restore the file to its original location with its original name.

Processes Tab Overview

The Processes tab is where the user assigns the data transformation Processes and uses them to link Input Fields with Output Fields. This linkage creates a data “flow” from the Input Fields (data “sources”) to the Output Fields (data “sinks”). Not all Input Fields need be used and some may be used more than once (linked to more than one Process). Output Fields are typically fully linked. Most Output Fields derive from SBDD geodatabase requirements (SBDD-source), but some have been added by the DIVA design team for their utility based on past experience (DIVA-source). Some SBDD-source Output Fields will have a “MustHaveData” Rule created for them if the Output Field is required to have data (per the SBDD geodatabase).

Note that only [Column Processes](#) are placed and manipulated on the Process Mapping Canvas. There are eight DIVA-defined Column Processes, ranging from a simple copy to complex decompositions. [Row Processes](#) (of which, at this time, there is only one) will automatically appear under the Output Fields List as needed. The existence of a Row Process is determined by the existence of the corresponding [Row Rule](#) (the user cannot create or remove this Process directly); the Process is then used to resolve violations detected by the corresponding Rule.

At DIVA start-up, the Processes tab will be inaccessible. Users must select an existing Submittal that has loaded data, or define a new Submittal and load its data, before the Processes tab will be available for use. In the image below are numeric keys to the discussion that follows. More detailed information on each Process type is referenced in the hyperlinks.



Key	Control	Description
1	Input Fields list	The list of Input Fields created when the active Input File was loaded. These fields are Data Sources .
2	Process Mapping canvas	This is a drag-and-drop area where users place and connect Processes with Input Fields and Output Fields or with other Processes through Process Fields.
3	Output Fields list	The list of Output Fields as imported from the SBDD geodatabase for the user-selected Period-Type (Reporting Period and Submittal Type). These fields are Data Sinks .
4	Record Pointer controls	Moves the record pointer on the Input File to (in order) the first, previous, next or last record.

Key	Control	Description
5	Process Palette	<p>Displays the graphics representing the eight data transformation processes defined in DIVA and serves as the “launch pad” for creating a Process. To create a Process, click one of the Process icons in the Palette, and drag-and-drop it onto the Process Mapping Canvas. The name of each Process can be found by hovering the mouse over its icon to display a tool tip. In order from left to right, these eight processes are listed below, with details on each found by following the hyperlink.</p> <ul style="list-style-type: none"> • Constant Value – populate an Output Field or another Process with a user-defined value. • Provider Value – populate an Output Field or another Process using values imported into DIVA from a provider information table. See “Update Provider Information” for details on this table. • Copy – populate the Output Field or another Process by copying the value of the Input Field. • Copy with Default – populate the Output Field or another Process where values are either null or blank by copying the default value as defined in the SBDD database for the active Reporting Period. • Make Negative – populate the Output Field or another Process by copying the value from the Input Field and inverting its sign. • Translation – populate the Output Field or another Process by translating values in the Input Field to a new, user-defined value. • Composition – populate the Output Field or another Process with values appended (composed) from two or more Input Fields. • Decomposition – populate two or more Output Fields or Processes with values parsed (decomposed) from a single Input Field. DIVA currently has four types of decomposition processes: Census Block, Full Address, Simple and Zip Code. The user cannot create a Decomposition Process, per se, but only one of these four types. While each is configured and functions in its own manner, the same symbol is used on all decomposition graphics. Associated tool tip can be used to identify specifics about these Processes.
6	Apply button	Applies the Processes, transforming Input Field values to Output Field values and causing them to display in red font below each Output Field.
7	Save button	Saves the Processes to a Processing Set. If already saved, this updates that Processing Set with any changes made since the last save.
8	Save As Template button	Saves the current Processing Set as a Template Processing Set for use when creating new Submittals.
9	Row Process(es) button(s)	<p>Below the Output Fields List, a button will be displayed for each existing Row Process. Click this button to open a dialog window to configure that Row Process.</p> <p><i>Note: At present, DIVA has only one defined Row Process.</i></p>

Tips for creating Processes:

- Preserve the Input Field, “InpSeqNo”, in the DIVA-defined “SeqNo” Output Field. This creates a unique record identifier which is useful in subsequent processing steps. Alternatively, use some other Input Field that uniquely identifies a data record (e.g., the UNQID field suggested for shapefile processing). Failing to put a unique identifier in the “SeqNo” Output Field can make it difficult to track NTIA-submitted data back to its source.
- Apply and save Processes as you define them. This ensures work is tested and then saved. Always examine the results of applying Processes across multiple data records to ensure that you have achieved what was intended.
- Use a “Provider Value” process instead of copying a Provider’s name, DBA name, FRN, etc. information from the Input File. This keeps the Provider identification information consistent with the values maintained in the Provider Information database and consistent across multiple Reporting Periods.
- Use Template Processing Sets when loading a new Submittal, or Copy and Paste another Submittal’s Processing Set. Chances are the processing will be similar enough that it will be faster to start with some pre-defined Processes.
- Right clicking any Process on the Process Mapping Canvas brings up a context menu which allows the user to change the configuration of inputs and/or outputs to the Process. Only composition allows you to add inputs, only decomposition denies adding outputs. All Processes allow you to delete them. Deletion can also be accomplished by selecting the input or output on the Process and pressing the “Del” (delete) key.

Rules Tab Overview

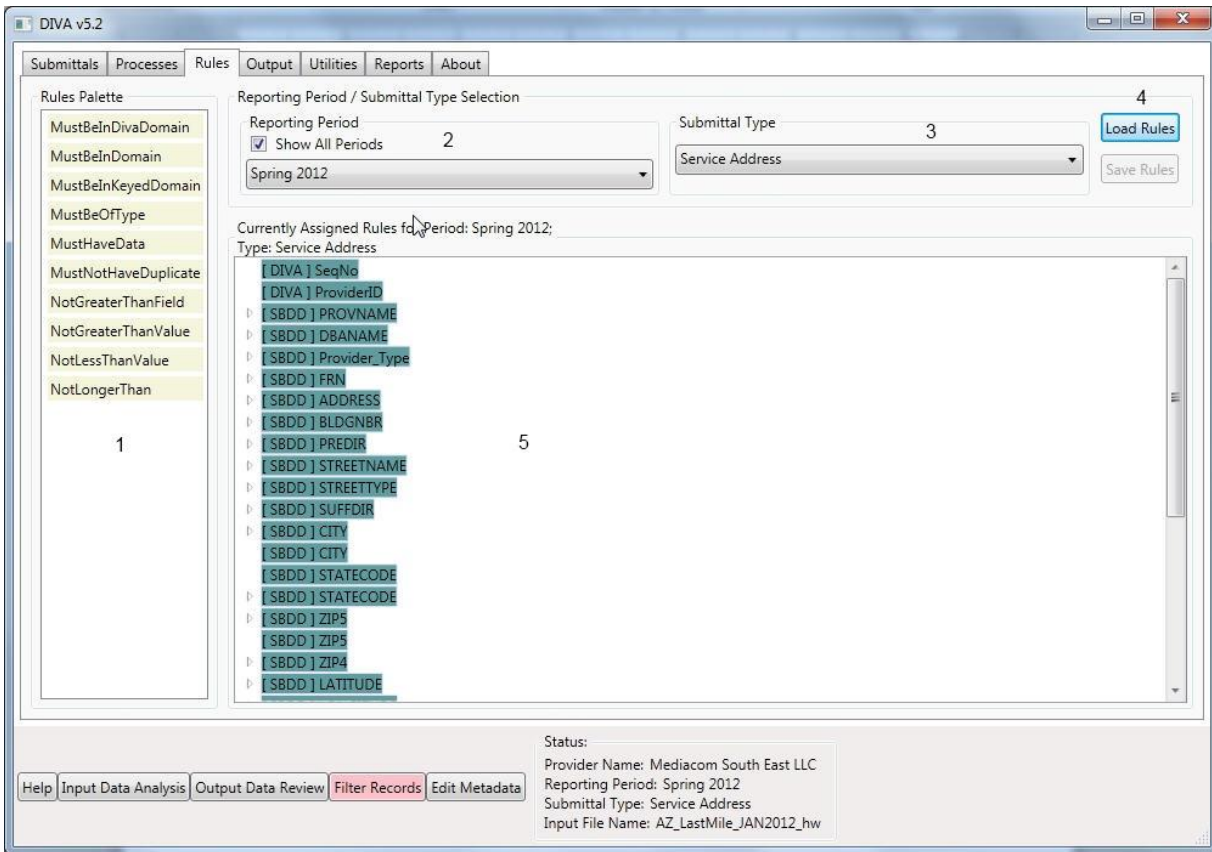
The Rules tab is where a user can review the Rules that are imported during the “SBDD XML Import” process. The user may also assign more Rules, modify (configure) existing Rules, or remove Rules. Rules are how DIVA implements [Data Verification](#) and [Data Validation](#) for Output Field values.

It is vital that the user understands that Rules are assigned to Template Output Fields. Template Output Fields are unique for each Period-Type (e.g., for Spring 2012 Census Block reporting). Each Input File that is identified with a Submittal (and there must be at least one) will have a corresponding set of Output Fields created; one Output Field for each Template Output Field. The Rules that are in effect for a Template Output Field affect every Output Field that is created from it. Because of this, all Submittals of a particular Period-Type will have the same set of Rules applied to each of them.

Obviously, this will not be true if the user processes some Submittals and then changes the Rules. Therefore, careful consideration should be given at the start of the Reporting Period, after SBDD geodatabase import, as to which Rules should be added or removed, and how they should be configured. **For consistency, when Rules are modified, previously processed Provider data should be re-processed and checked for consistency with these modified Rules.** DIVA makes it easy to rapidly review any previously processed data.

Because Rules are assigned to Template Output Fields, it is not necessary to have an active (selected) Submittal, or even any defined Submittals for a Reporting Period, in order to access and modify the [Rule Set](#) for a Period-Type. All that is required is that the SBDD geodatabase be imported for the Reporting Period. Nor is it required that a Rule Set be loaded on this tab before they are applied (the results of Rule application are seen in the Output Tab). They will automatically be applied because the Output Field is tied to the Template Output Field from which it was derived

In the image below are numeric keys to the discussion that follows. More detailed information can be found by following the hyperlinks.



Key	Control	Description
1	Rules Palette area	A list of rule types that can be applied by dragging and dropping them onto the Output Fields shown in the “Currently Assigned Rules” area.
2	Reporting Period pull-down list	This pull-down list displays the current Reporting Period or all of the Reporting Periods defined in DIVA if the “Show All Periods” box is checked.
3	Submittal Type pull-down list	Each of the eight NTIA approved Submittal Types are listed in this menu.
4	Load Rules button	Loads the Rule set for the selected Period-Type.
5	Currently Assigned Rules area	Displays the Rules that have been assigned to each Template Output Field of a Period-Type (and, thus, each Output Field of a Submittal).

Output Tab Overview

The Output tab is where a user reviews and evaluates the results of applying Rules to values in the Output Fields. Users click the “Refresh Violation Counts” button to update the violation counts (apply the Rules). Rules and their violation counts are shown in a tree-view grouped under the Output Field to which they apply. Double-clicking on any Output Field will toggle the expansion / contraction of the tree view. Double-clicking on any given Rule that shows an error displays details on that error in the “Failed Record View”. Depending on the Rule, additional information may be displayed in the Extra Data panel below the “Error Record View” area.

In the image below are numeric keys to the discussion that follows. More detailed information on each entry can be found by following the hyperlink in the table.

Output Actions

Refresh Violation Counts 1

Rule Violations and Counts

2

Failed Record View for FULLFIPSID => Must Not Have Duplicate Records with Field TRANSTECH 3

InpSeqNo	FULLFIPSID	TRANSTECH	MAXADDOWN	MAXADUP	TYPICDOWN	TYPICUP
304	040030011003032	40	7	4	6	2
305	040030011003032	40	6	4	7	4
254	040030011003047	40	7	4	7	3
255	040030011003047	40	7	3	7	4
256	040030011003053	40	5	4	7	5
257	040030011003053	40	6	2	7	7
258	040030011003053	40	7	4	7	
259	040030011003053	40	5	3	7	4
260	040030011003059	40	7		7	4
261	040030011003059	40	4	5	7	4
262	040030011003059	40	5	4	7	3
263	040030011003059	40	6	2	7	4
264	040030011003059	40	7	3	7	
301	040030012002012	40	7	4	5	4
302	040030012002012	40	6	3	6	2
303	040030012002012	40	7	4	7	3
306	040030012002019	40	7	3	7	3
307	040030012002019	40	6	3	6	3
308	040030012002019	40	5	4	4	4

Duplicate rows as output 4

InpSeqNo	FULLFIPSID	TRANSTECH	MAXADDOWN	MAXADUP	TYPICDOWN	TYPICUP
304	040030011003032	40	8	8	9	4
254	040030011003047	40	7	4	7	4
255	040030011003053	40	7	4	7	

Status:

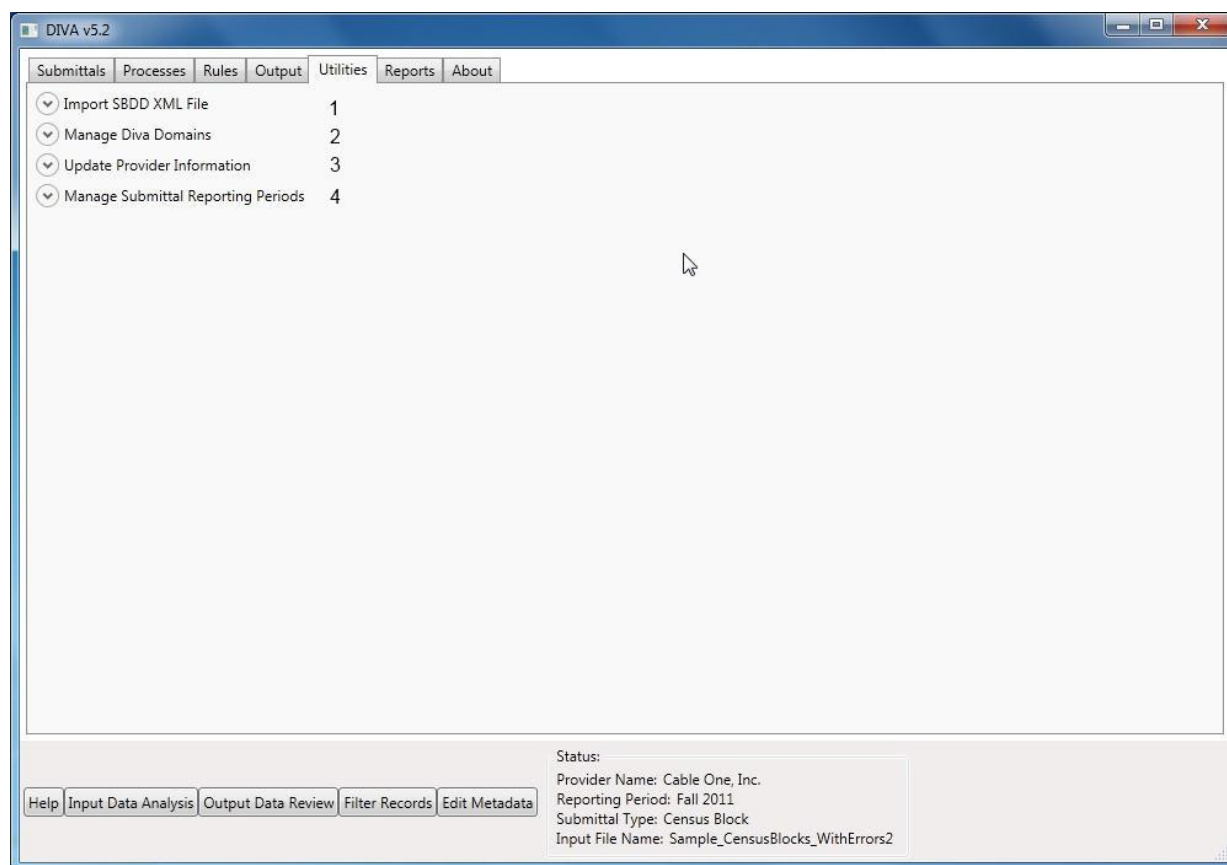
Provider Name: Cable One, Inc.
Reporting Period: Fall 2011
Submittal Type: Census Block
Input File Name: Sample_CensusBlocks_WithErrors2

Help Input Data Analysis Output Data Review Filter Records Edit Metadata

Key	Control	Description
1	Refresh Violations Count button	Click this button to update the record violation counts in the “Rules and Violation Counts” area. Be sure Processes have been applied prior to refreshing violation counts by clicking on the “Processes” tab first.
2	Rules and Violation Counts area	<p>Displays the results of applying Rules to the Output Fields of the active Submittal. Any number greater than or equal to zero is the error count. A “-2” indicates that the Rule is not configured and therefore cannot be evaluated. A “-1” indicates that there is no data in the Output Field to be evaluated. Either the Output Field is not connected to a data source on the Processes tab or there is no data in the connected data source.</p> <p><i>Note: There is a “splitter” line between the Output Actions / Rule Violations and Counts panels and the Failed Record View / Extra Data panels. The user may click and drag this line to change the panel area sizes in the window.</i></p>
3	Error Record View	Displays rows which have an error in the Output Field for the Rule that has been selected and double-clicked in the “Rules and Violation Counts” area.
4	Duplicate Rows as Output area	<p>This area is used to display either details about the “MustNotHaveDuplicate” error type or the Domain values for domain rule types.</p> <p>For duplicate rules It displays the record(s) that will be added after compositing the duplicates into a single record based on user-defined parameters.</p> <p><i>Note: There is a “splitter” line between the Failed Record View and the Extra Data areas. The user may click and drag this line to change the panel area sizes in the window.</i></p>
4	Domain Table Display Area	<p>This area is used to display either details about the “MustNotHaveDuplicate” error type or the Domain values for domain rule types.</p> <p>For domain rule types, it will display the domain table when the user clicks on a domain Rule.</p> <p><i>Note: There is a “splitter” line between the Failed Record View and the Extra Data areas. The user may click and drag this line to change the panel area sizes in the window.</i></p>

Utilities Tab Overview

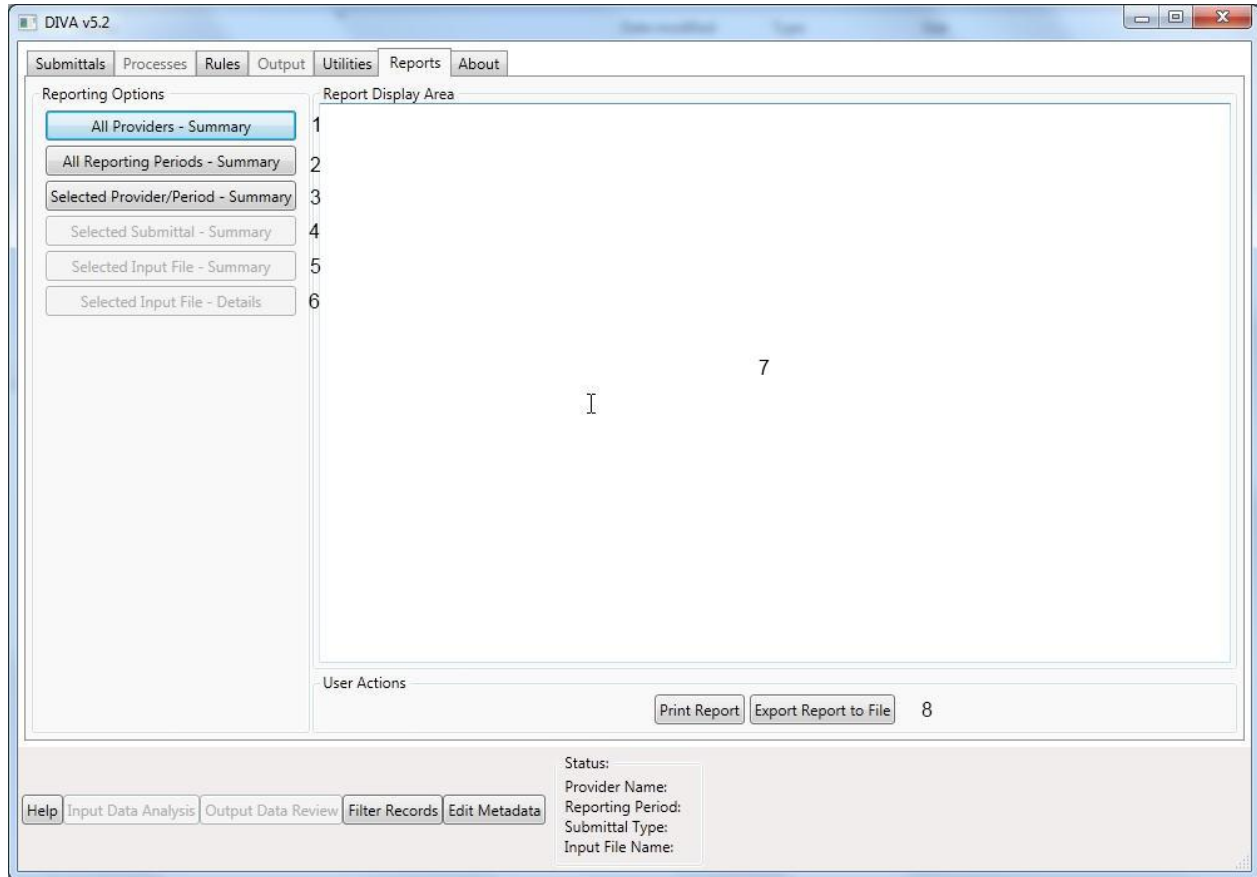
The Utilities tab is where a user configures various aspects of DIVA to keep it current with new information. In the image below are numeric keys to the discussion that follows. More detailed information on each utility can be found by following the hyperlinks.



Key	Control	Description
1	Import SBDD XML File	Import an XML workspace document that describes an SBDD geodatabase schema and domain tables. This need only be done once for each Reporting Period.
2	Manage DIVA Domains	Import a text, Excel or DBF containing DIVA Domain values.
3	Update Provider Information	Import a text, Excel or DBF containing Provider identification and status information. The field names in the source file must match exactly with those that DIVA expects.
4	Manage Submittal Reporting Periods	Add or remove Reporting Periods, or remove all data for a Reporting Period.

Reports Tab Overview

The Reports tab is where a user selects from a menu of pre-configured reports and either prints or exports them. Note that in the illustration below, some report options are greyed out. This is because no Submittal has been selected on the Submittals tab. More detailed information on each report and control can be found by following the hyperlinks in the summary table.



Key	Control	Description
1	All Providers Summary button	Displays a list of information about each Provider currently loaded into DIVA. Use the “Update Provider Information” button on the “Utilities” tab to update this information.
2	All Reporting Periods Summary button	Displays a list of information about each Reporting Period currently defined in DIVA. Use the “Manage Submittal Reporting Periods” button on the “Utilities” tab to update this information. This report will include all notes recorded by the user on the left-hand side of the “Edit Metadata” dialog.

Key	Control	Description
3	Selected Provider/Period Summary button	<p>Displays a list of information about the active Provider and Reporting Period, including a list of the Period-Types (Submittal Types) defined for the Reporting Period and the number of existing Submittals for that Period-Type.</p> <p><i>Note that the Submittal count can only be 0 (none defined) or 1 (a Submittal exists). For brevity, only Period-Types that have a Submittal count of 1 are displayed.</i></p>
4	Selected Submittal – Summary button	Displays a summary of the active (selected) Submittal, including summary information about each of its Input Files (inaccessible if no active Submittal).
5	Selected Input File – Summary button	<p>Displays more detailed information about the active (selected) Input File of the active (selected) Submittal on the Submittals tab(inaccessible if no active Submittal).</p> <p>Information will be displayed about each Output Field of the Input File followed by a “chain of data flow” that shows the Output Field’s connected Process, any Process Fields connected to that Process, any Processes connected to those Process Fields, etc., until the ultimate Data Source (which may be a Process or an Input Field) is listed.</p> <p>After the “chain of data flow” into the Output Field, a list of the Rules that apply to this Output Field is displayed along with the number of current Rule violations.</p>
6	Selected Input File – Details button	<p>Same as “Selected Input File – Summary” – except that additional details are given.</p> <p>Some Processes will have additional information given, e.g., for the Translation Process, all translated From and To values will be displayed.</p> <p>For Rule listings that have a violation count greater than 0, all violating records will be listed (for brevity, this includes just the record number [“InpSeqNo” field] and the Output Field’s value). This report can be very long if there are excessive Rule violations, so the Summary version should be checked before generating this report.</p>
7	Report Display area	The area in which the various reports are displayed. Scroll bars will automatically appear if necessary.
8	User Actions area	Allows the user to either print the file to a printer or export it to a text file.
9	Print Report button	Prints the information in the Report Display area to a printer selected by the user (invokes a standard Window’s Printer dialog box).

Key	Control	Description
10	Export Report to File button	Exports the information in the Report Display area to a file, in plain text format, as selected by the user (invokes a standard Window's File Save dialog box).

About Tab Overview

The About tab gives general information about DIVA, its current version number, contact information, acknowledgements for other code uses and the licensing information for that code. There is no other functionality for this tab.

Functionality Details

This section describes in detail how to perform key operations in DIVA.

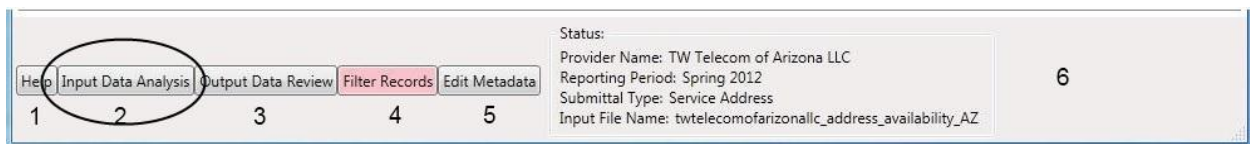
Status Bar Functions

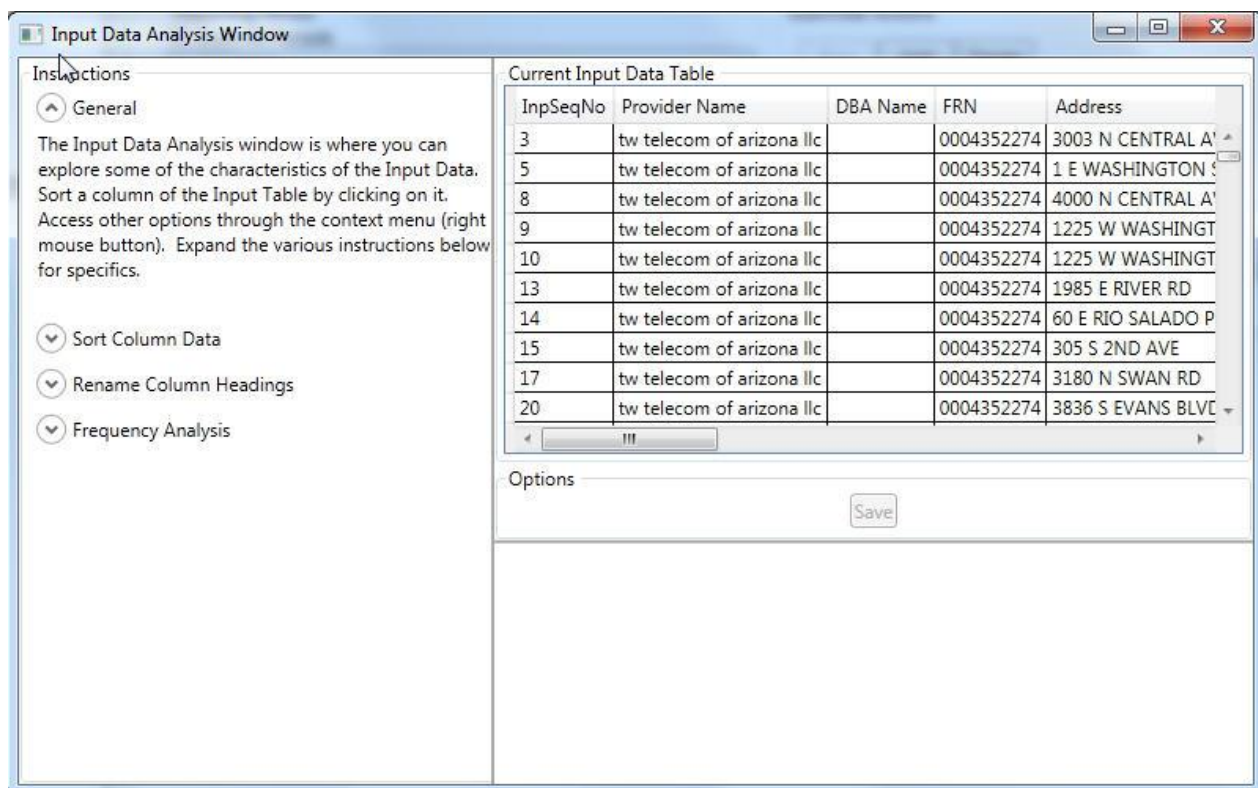
Input Data Analysis button (Status Bar, “Input Data Analysis” button) – the Input Data Analysis window shows the contents of the active Input File for the active Submittal. It offers a number of features for evaluating the data that has been loaded, including sorting and frequency analysis. It also has its own Help window which offers guidance on the use of its features.

Note: a Submittal must have its data loaded (not just created) before the Input Data Analysis Window can be accessed.

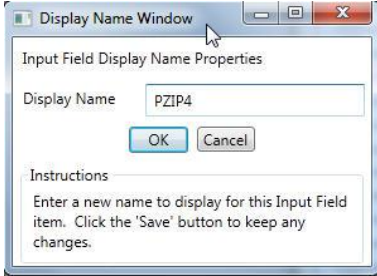
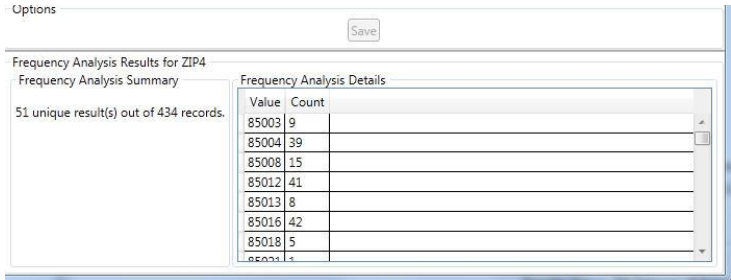
Note: This window is not a “modal” window; it may be left open while other operations are performed in DIVA. The user may minimize the window or may click the “Input Data Analysis” button to “toggle” this window in and out of view.

Note: The user must explicitly close this window, by clicking the red “X” icon in the window’s upper right, before selecting another Submittal. If it is left open, it will continue to display the originally selected Submittal’s data, not the currently active Submittal’s data.





Key	Control	Description
1	Instruction Panel	<p>Instructions to guide users on how to understand and use the Input Data Analysis Window.</p> <p>Note: there is a "splitter" line between the Instructions and the rest of the window which can be dragged to change the amount of the window allocated to each panel. Dragging the line all the way to the left will completely hide the instructions, but the splitter line is still there so that the instructions may be redisplayed at any time.</p>

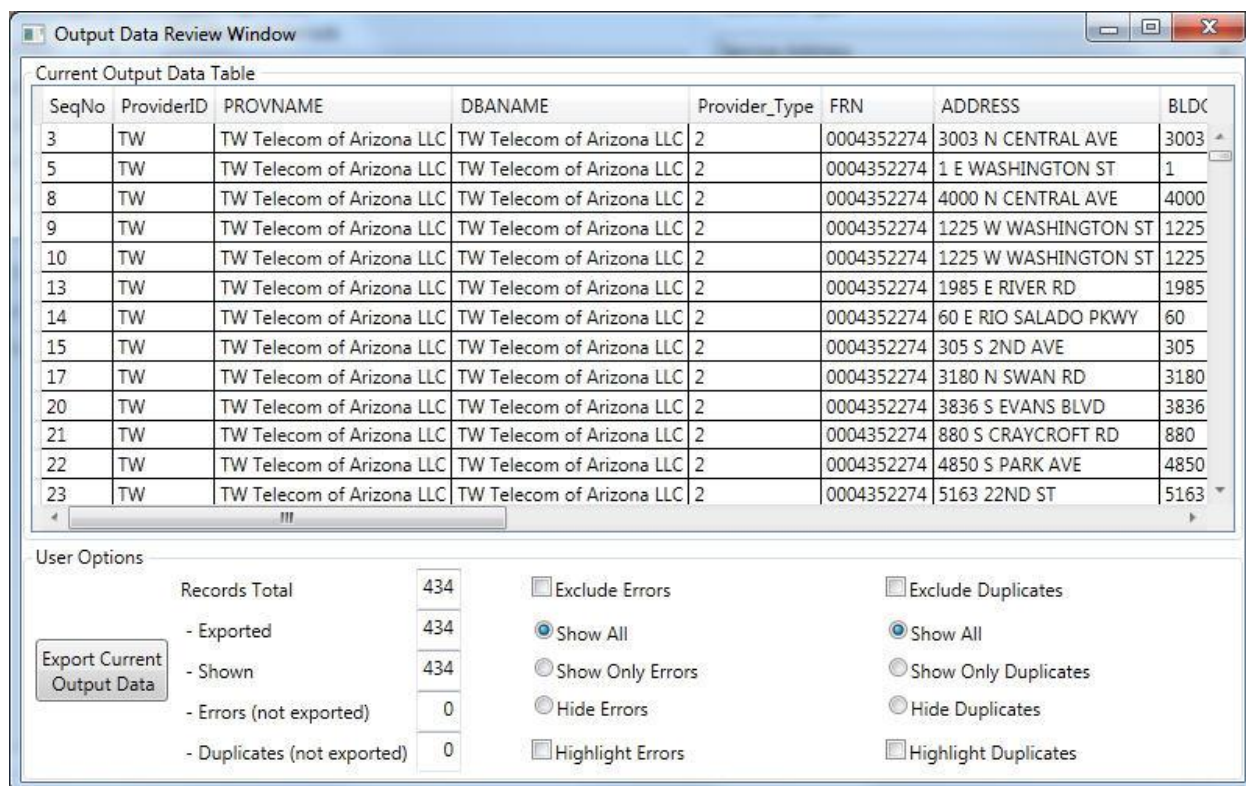
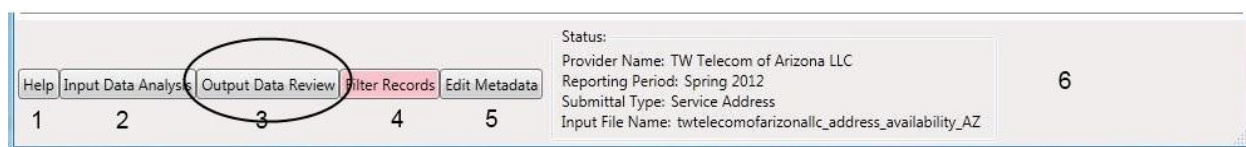
Key	Control	Description
2	Header Row Context Menu, “Rename the Header”	<p>Right Click on any column header and select “Rename the Header”</p> <p>Type in a new Display Name and click “OK” to save changes and dismiss the dialog. Click “Cancel” to exit without saving.</p>  <p><i>Note: You can use spaces in the Display Name. This string is just an alias and does not become an actual database field name.</i></p> <p><i>Note: Upon clicking “OK” which dismisses the dialog and returns you the Input Data Analysis Window, you must click the “Save” button immediately below the table display in order to have Display Name changes be saved permanently (i.e. so you will see them the next time you select the Submittal).</i></p>
2	Header Row Context Menu, “Do Frequency Analysis”	<p>Right Click on any column and select “Do Frequency Analysis” option</p> <p>The results of the Frequency Analysis are shown in its display area. On the left side, a summary of the number of unique values for the analyzed field (column) and total number of records in the Input File is displayed. On the right side, a detailed table of the analysis results is shown, with each unique value listed along with the frequency count.</p>  <p><i>Note: There is a “splitter” line between the Data Table and the Frequency Analysis Results areas. The user may click and drag this line to change the viewing area sizes</i></p>

Key	Control	Description
		<p><i>in the window.</i></p> <p><i>Note: Clicking on the header column of the results table will sort the entries. Thus, by clicking on the Count header, you can sort the results by the Count value in ascending order. Clicking again will sort the values in descending order.</i></p>
3	Save button	Saves any changes made to the header row (column) names.
4	Display area	The area where the “Do Frequency Analysis” output (table) is shown when that menu option is selected.
5	Cancel Menu control	<p>Click on the red x in the upper right corner of the “input Data Analysis Window” to dismiss it.</p> <p><i>Note: no matter what you do in this window, the source Input File is not changed.</i></p>

Output Data Review button (Status Bar, “Output Data Review” button) – the “Output Data Review” button opens the “Output Data Review Window” which shows the results of applying [Processes](#) to the values in the Input Fields and [Rules](#) to the values in the Output Fields of a Submittal. This window offers information about the processed file and options to apply before exporting it to a text file.

Note: This window may be left open while other operations are performed in DIVA. The user may minimize the window or may click the “Output Data Review” button to “toggle” this window in and out of view.

Note: The user must explicitly close this window, by clicking the red “X” icon in the window’s upper right, before selecting another Submittal. If it is left open, it will continue to display the originally selected Submittal’s data, not the currently active Submittal’s data.

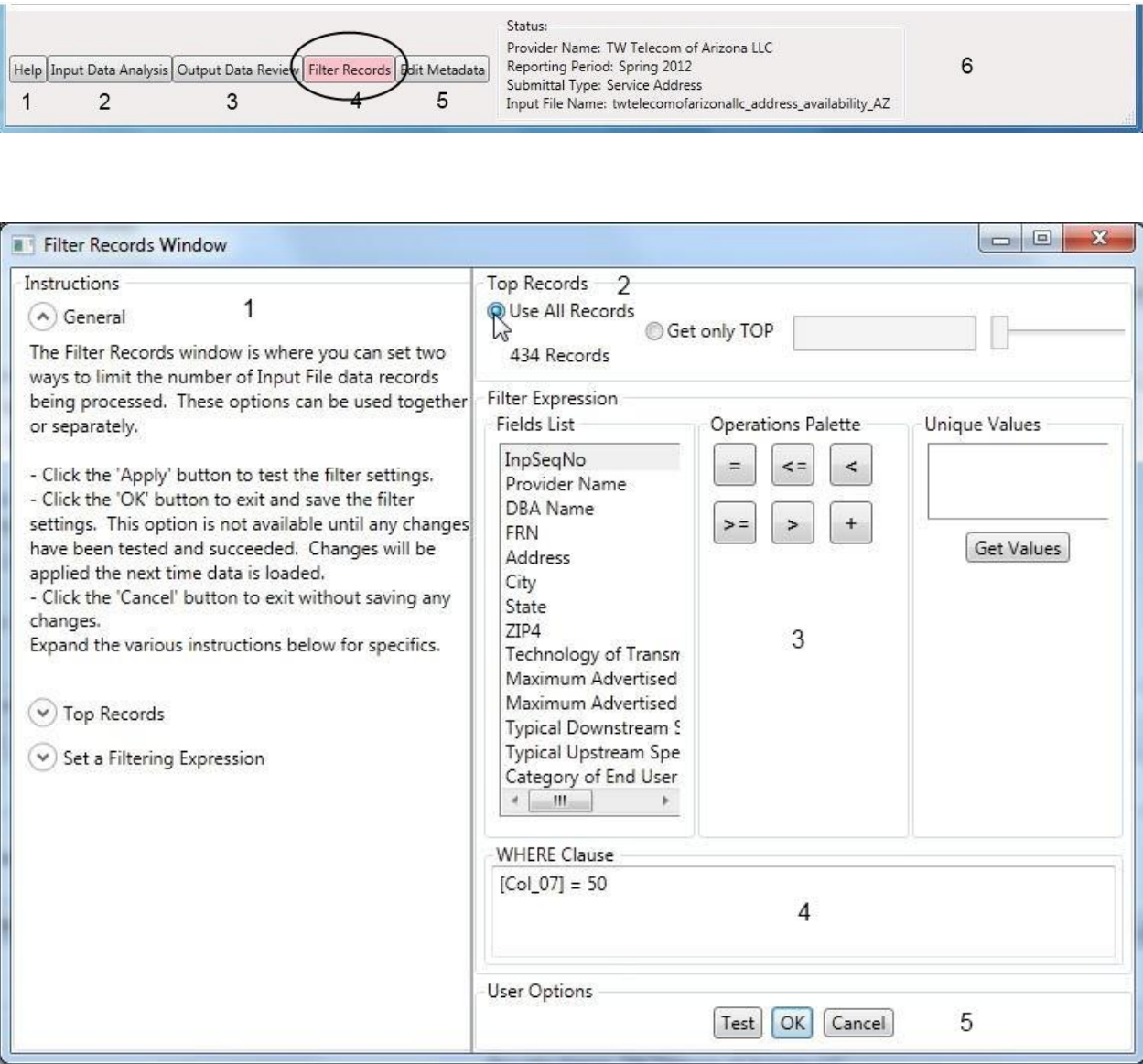


Key	Control	Description
1	Export Current Output Data button	Opens a dialog box where the user can specify the output location and file name for data that will be exported. Export only works on records visible in the “Current Output Data Table” control and only produces delimited text files.
2a	Records Total text box	Total records in the active Submittal. This number reflects any Filters the user has applied.
2b	Exported text box	The number of records that will be exported when the “Export Current Output Data” button is clicked.
2c	Shown text box	The number of records being displayed in the “Current Output Data Table” section of the window. This number reflects any errors or duplicates the user has chosen to exclude using the checkboxes in this area.
2d	Errors (not exported) text box	The number of records with an error that will not be exported. This number will differ from the Total Records, only if the user has checked the “Exclude Errors” box in the “Errors Column”.
2e	Duplicates (not exported) text box	The number of duplicate records that will not be exported. This number will differ from the Total Records, only if the user has checked the “Exclude Duplicate” box in the “Errors Column”.
3a	Exclude Errors checkbox	Check this box to exclude errors from the export file.
3b	Show All radio button	Show all records, including those with errors.
3c	Show Only Errors radio button	Show only records that have been flagged with an error.
3d	Hide Errors radio button	Show only records that have NOT been flagged with an error.
3e	Highlight Errors check box.	Check this box to highlight error rows so they can more easily be seen.
4a	Exclude	Check this box to exclude duplicate records from the export file.

Key	Control	Description
	Duplicates checkbox	
4b	Show All radio button	Show all records, including those with duplicates.
4c	Show Only Duplicates radio button	Show only records that have been flagged as duplicates.
4d	Hide Duplicates radio button	Show only records that have NOT been flagged as duplicates.
4e	Highlight Duplicates checkbox	Check this box to highlight duplicate rows so they can more easily be seen.

Note: Errors and duplicates are two separate concepts and mutually exclusive when displaying in this window. If a record is both a duplicate and an error, it will be flagged as an error first and will remain that way until the other rule violation(s) is/are fixed. Only then will its duplicate status be considered.

Filter Records Window (Status Bar, “Filter Records” button) – the “Filter Records” window allows users to define a subset of data contained in a Provider Input File for further processing. Filters affect how many records are loaded from a Provider’s Input File as well as any data that has been loaded into DIVA using the Load button on the Submittals Tab. Filters do not change the input data in any way.



Key	Control	Description
-----	---------	-------------

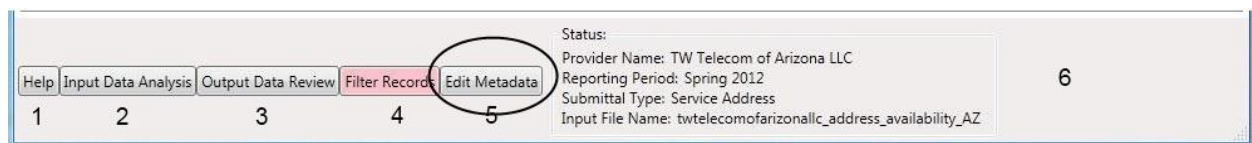
Key	Control	Description
1	Instruction panel	<p>Instructions to guide users on how to understand and use Filters.</p> <p>Note: there is a “splitter” line between the Instructions and the rest of the window which can be dragged to change the amount of the window allocated to each panel. Dragging the line all the way to the left will completely hide the instructions, but the splitter line is still there so that the instructions may be redisplayed at any time.</p>
2	Top Records Selection area	<p>Filter by selecting the top N rows per user specification.</p> <p>Click the “Use All Records” radio button to use the entire data set.</p> <p>Click the “Get only TOP” radio button to set the number of records to be retrieved.</p> <p>Type the number of records to be retrieved into the text entry box next to this radio button OR use the slider bar control to adjust the record number.</p> <p>Note that this Filter can be applied before data is initially loaded. If that is the case, then the number of rows that may be user-specified is between 4,000 and 40,000 (DIVA has no way of knowing how many records are in the Provider’s data set at this time, so this is completely arbitrary). Setting this filter before data loading does not affect loading all of the Provider’s data; it merely affects how many of those records will be initially available for processing. After data loading, the number of records is limited to 10% to 60% of the total number of records.</p> <p>Note: setting a filter (clicking “OK”) has no immediate effect on the Submittal data currently loaded into DIVA for processing. The user must switch to another Submittal (or Input File of the current Submittal, if there is one) and then back to this Submittal (or Input File) for the filtering to take effect.</p> <p>Note: close the Input Data Analysis and / or Output Data Review Windows before re-loading the data to have them properly reflect the change.</p>

Key	Control	Description
3	Filter Expression area	<p>Use the controls in this area to build an SQL “Where” Expression to filter records:</p> <ol style="list-style-type: none"> A list of fields from the active Input File is presented on the left side in the “Fields List”. Double-clicking a displayed field name will add it to the “WHERE Clause” expression building panel. An “Operations Palette” is presented in the middle from which a user can select operators for the expression (e.g. less than). Clicking one of the buttons will add that operation to the “WHERE Clause” expression at the current insertion point. A “Unique Values” box is shown on the far right. A user may wish to see the unique values present in the field chosen from the “Fields List” to help build an expression. Clicking the “Get Values” button will retrieve a (sorted) list of unique values for the currently selected Field in the Fields List. The “WHERE Clause” area (see #4 below) shows the query expression as it is being built.
4	Where Clause area	<p>This is the “Where Clause expression building panel. It is where the user can see the query expression being built. It is an area in which the user can also directly edit the query expression.</p>
5	User Options	<p>The “User Options” area is where the “WHERE Clause” filter expression is tested (“Test” button) and then accepted/applied (“OK” button) or canceled (“Cancel” button). Expressions must be tested before they can be applied (the “OK” button will not be activated until the expression is successfully tested).</p>

Edit Metadata button (Status Bar, “Edit Metadata” button) – the Edit Metadata window provides a place for users to record notes about the active Provider (general comment), the active Reporting Period (general comment), the active Provider for the active Reporting Period (general comment), the active Period-Type (determined by the active Reporting Period / active Submittal Type combination; general comment), the active Submittal (specific comment) and / or the active Input File for the currently active Submittal (numerous specific comments).

To enter a comment about a Provider, for example, that Provider must be the active (selected) Provider. Thus the metadata targets in this window are dependent on the context of what is currently selected in the Submittals Tab. If there is no active Submittal, then the Submittal and Input File entry panels will be inaccessible.

Users can choose to save their work or cancel without saving any changes. The information entered into this form is available for viewing by re-opening this window or through one or more reports on the “Reports” tabbed dialog.



The screenshot shows the 'Edit Metadata Window' with the following components and numbered callouts:

- 1**: Notes for Provider: TW Telecom of Arizona LLC (text box)
- 2**: Notes for Reporting Period: Spring 2012 (text box)
- 3**: Notes for Provider: TW Telecom of Arizona LLC for Reporting Period: Spring 2012 (text box)
- 4**: Notes for Reporting Period: Spring 2012 Submittal Type: Service Address (text box)
- 5**: Notes for selected Submittal: ID=2, Provider: TW Telecom of Arizona LLC, Period: Spring 2012, Type: Service Address. No FULLFIPS ID: will need to add through GIS overlay process (text box)
- 6**: Input File Notes: twtelecomofarizonallc_address_availability_AZ (text box)
- 7**: File Name Comments: (text box)
- 8**: Projection Type: (text box)
- 9**: Projection Comments: (text box)
- 10**: Header Row Comments: (text box)
- 11**: User actions (Save Notes, Cancel Changes buttons)

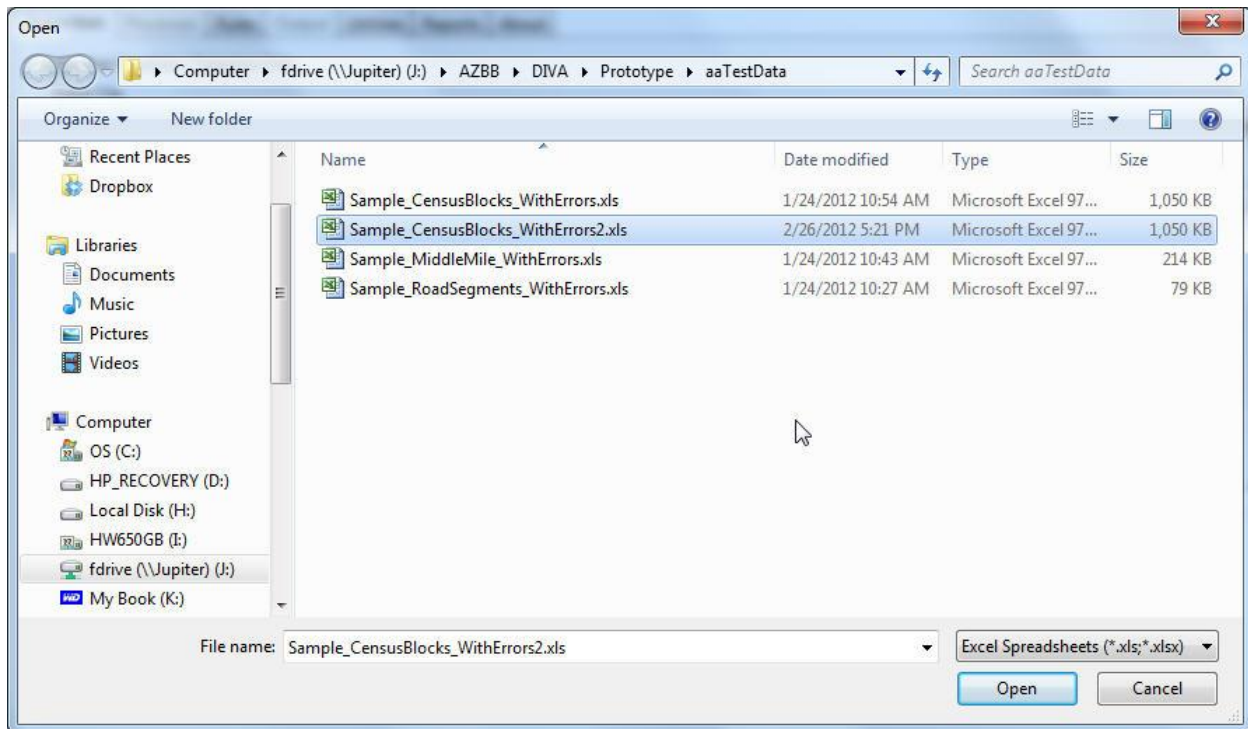
Key	Control	Description
1	Notes for Provider text box	<p>Enter notes pertaining to the Provider (not specific to any Reporting Period or Submittal Type). For example, "Provider X generally submits data in CAD format and we need to contact them ahead of time to ask for shapefiles."</p> <p>A comment in the Provider metadata panel will show up in all reports for that Provider, no matter what the Reporting Period is or what Submittal is selected for that Provider.</p>
2	Notes for Reporting Period text box	<p>Enter notes pertaining to the Reporting Period (not specific to any Provider or Submittal Type). For example, "In the Spring 2012 Reporting Period we began to send Verification Packages to all Providers even if they did not submit for that Period. We re-sent them a package from a previous Reporting Period."</p>

Key	Control	Description
3	Notes for Provider for Reporting Period text box	<p>Enter notes for the selected Provider for the selected Reporting Period (not specific to any Submittal Type). Use this panel to record metadata that applies to a Provider for all data submitted during the Reporting Period. This is a good place to record general questions to follow-up with the Provider, as well as overall processing notes.</p> <p>A comment in this panel will show up in all Submittal and / or Input File related reports for that Provider / Reporting Period combination.</p>
4	Notes for Reporting Period and Submittal Type text box	Use the Period-Type panel to record SBDD geodatabase-related metadata as this information will show up for every Submittal of that Period-Type for every Provider. This is a good place to record additional Rules that have been added for that Period-Type and why they were added.
5	Notes for selected Submittal text box	Enter notes specific to the active Submittal. For example, "No FULLFIPSID; we will have to obtain through overlay process."
6	Notes for Input File text box	Enter notes specific to the Input File associated with the selected Submittal. For example, "The source file was not in column-row format and had to be transformed by TSSW using Excel."
7	File Name Comments text box	Enter notes specific to the Input File's data source file name. For example, "File Name indicates a Census Block delivery, but it really is an address submittal as well."
8	Projection Type text box	Enter the projection. This only applies to shapefile submittals.
9	Projection Comments text box	Enter any comments about the projection. For example, "The Census Block shapefile was delivered in WGS 84 while the Roads were in UTM Zone 12."
10	Header Row Comments text box	Enter any comments about the header row in the Provider delivery. For example, "No header row; TSSW had to guess as to which column mapped to the proper SBDD field."
11	User Actions buttons	Click "Save Notes" to exit and save your changes. Click "Cancel Changes" to exit <u>without</u> saving edits.

Submittals Tab Functions

Select an Input File (Submittals Tab, “Input File” control, “Browse” button) – the “Browse” button opens a file-browser dialog which lets the user navigate to any area on their network to select an Input File for further processing.

Click on the Browse button located on the Input File control of the Submittals Tab and the Open file dialog appears



Key	Control	Description
1	Folder pane	Use the folder pane to browse to the folder on your system or network where the Provider source file is located.
2	File pane	Use the file pane to select the the Provider source file.

3	Filename text box	Displays the name of the selected file.
4	File Type Pull-down	Set the File Type pull-down to match the type of Input File for which you are browsing. Only delimited text, DBF and Excel file types are supported by DIVA.
5	User Option buttons	Click the “Open” button after selecting the desired file or double-click on the desired file to select it. The file-browse dialog is dismissed and the user is returned to the DIVA submittals tab, with the path and file name now showing in the Input File text box. Click the “Cancel” button to dismiss dialog without selecting a file.

Special Note on Processing Shapefiles

DIVA is not designed to process shapefiles. However, users can load the DBF portion of a shapefile, transform its data, and then join the output text file back to the original shapefile. Users cannot change record count from the source or the join operation will not be valid. See “[Processing Shapefiles](#)” for further details.

Preview an Input File (Submittals Tab, “Input File” control. “Preview” button) – the Preview Window is a common control called from a number of different places in DIVA. This is simply a preview function and neither affects the source file being previewed, nor loads the source data into DIVA. Its purpose is to let the user preview a source file prior to loading it into a DIVA table. If major problems are encountered, the user can then take appropriate steps to modify the source file, filter the number of in-coming records or to perform other tasks to ensure the Input File will be loaded into DIVA cleanly.

Click on the Preview button located on the Input File control of the Submittals Tab and the “Preview Window” opens. (Note: for performance reasons, this window only displays the first 100 records of the Provider’s data set. This setting is not user-controllable).

Provider Name	DBA Name	FRN	Address	City	State	ZIP4	Technology
tw telecom of arizona llc		0004352274	4400 E BROADWAY BLVD	TUCSON	AZ	85711	30
tw telecom of arizona llc		0004352274	15601 N 40TH ST	PHOENIX	AZ	85032	30
tw telecom of arizona llc		0004352274	3003 N CENTRAL AVE	PHOENIX	AZ	85012	50
tw telecom of arizona llc		0004352274	1055 N LA CANADA DR	GREEN VALLEY	AZ	85614	30
tw telecom of arizona llc		0004352274	1 E WASHINGTON ST	PHOENIX	AZ	85004	50
tw telecom of arizona llc		0004352274	1955 S VAL VISTA DR	MESA	AZ	85204	30
tw telecom of arizona llc		0004352274	2928 S 38TH ST	PHOENIX	AZ	85040	30
tw telecom of arizona llc		0004352274	4000 N CENTRAL AVE	PHOENIX	AZ	85012	50
tw telecom of arizona llc		0004352274	1225 W WASHINGTON ST	TEMPE	AZ	85281	50
tw telecom of arizona llc		0004352274	1225 W WASHINGTON ST	TEMPE	AZ	85281	50
tw telecom of arizona llc		0004352274	4040 E CAMELBACK RD	PHOENIX	AZ	85018	30
tw telecom of arizona llc		0004352274	15353 N 91ST AVE	PEORIA	AZ	85381	30
tw telecom of arizona llc		0004352274	1985 E RIVER RD	TUCSON	AZ	85718	50
tw telecom of arizona llc		0004352274	60 E RIO SALADO PKWY	TEMPE	AZ	85281	50
tw telecom of arizona llc		0004352274	205 S 7TH AVE	PHOENIX	AZ	85002	50

Key	Control	Description
1	File Type pull-down	Set the File Type pull-down to match the type of Input File for which you are browsing. Only delimited text, DBF and Excel file types are supported by DIVA.
2	Delimiter pull-down	If you are opening a delimited text file, then you must select the appropriate delimiter from this pull-down menu. When opening DBF or Excel files this option has no effect.

Key	Control	Description
3	First Row has Header checkbox	Click the “First Row Has Headers”, toggling the checkbox to match the structure of your Input File regarding the presence of a header row. If your file does not have a header row (column names), and this box is left unchecked, DIVA will name the columns in sequency (e.g., “Col01”, “Col02”, etc.)
4	Refresh button	Click this button to load the file or refresh that file display after making changes.
5	Column Headings	Double-click any column header to toggle the table sorting in ascending or descending order of values in that column.
6	Table Display area	This area displays the first 100 records of the source file. This display is a preview of how the file will look when you actually load the data.
7	Dismiss Window	The red “X” in the upper right corner is used to dismiss the “Preview Window”

Load an Input File (Submittals Tab, “Load Button”) – the Load button copies in records from the current Input File into an internal DIVA table. This step is required before any further processing can be accomplished.

Selected Submittal's Input File Information

Sample_RoadSegments_WithErrors

Delete Load

Processing Set

Copy Paste

Click on the Load button on the Submittals tab and the Load Window opens. This dialog is very similar to the “Preview Window” previously discussed, but with the option to load the data into DIVA.

Load Window

File Type: Excel Delimiter: Tab ☒ First row has Headers Load Refresh

ProName	DBAName	FRN	CenBlock	TechTrans	MaxAdDnSp	MaxAdUpSp	TypDnSp	TypUpSp	E
Cable One	Cable One	0003474327	040030010001032	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001045	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001046	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001048	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001049	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001050	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001052	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001055	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001060	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001061	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001062	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001063	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001066	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001069	41	4	7	7	4	5
Cable One	Cable One	0003474327	040030010001070	41	4	7	7	4	5

Key	Control	Description
1	File Type pull-down	Set the File Type pull-down to match the type of Input File for which you are browsing. Only delimited text, DBF and Excel file types are supported by DIVA.
2	Delimiter pull-down	If you are opening a delimited text file, then you must select the appropriate delimiter from this pull-down menu. When opening DBF or Excel files this option has no effect.

Key	Control	Description
3	First Row has Header checkbox	Click the “First Row Has Headers”, toggling the checkbox to match the structure of your Input File regarding the presence of a header row. If your file does not have a header row (column names), and this box is left unchecked, DIVA will name the columns in sequence (e.g., “Col01”, “Col02”, etc.)
4	Load button	Click this button to load the data into DIVA.
5	Refresh button	Click this button to load the file or refresh that file display after making changes.
6	Column Headings	Double-click any column header to toggle the table sorting in ascending or descending order of values in that column.
7	Table Display area	This area displays the first 100 records of the source file. This display is a preview of how the file will look when you actually load the data.
8	Dismiss Window	The red “X” in the upper right corner is used to dismiss the “Preview Window”

After the data is loaded you will see the name assigned to the Input File’s Processing Set in the “Processing Set” area of the Submittals tab.

DIVA v5.2

Submittals | Processes | Rules | Output | Utilities | Reports | About

Submittal Creation / Selection Info

Input File: Browse Clear Preview

Provider: Name: Cable One, Inc. DBA: Cable One FRN: 0003474327

Submittal Type: Census Block

Template ProcessingSet: ☐ Use current Template Processing Set

Reporting Period: ☒ Show All Periods Fall 2011

Submittal Actions: New Add Delete

Existing Submittals: Selected Submittal is Target of all Editing

Filter Submittals by: ☐ Providers ☐ Periods ☐ Types

Submittal ID	Provider ID	Provider Name	Submittal Period	Submittal Type
8	C1	Cable One, Inc.	Fall 2011	Census Block
9	CC	Comcast of Arizona Inc.	Fall 2011	Road Segment

Selected Submittal's Input File Information

Sample_CensusBlocks_WithErrors2 Delete Load

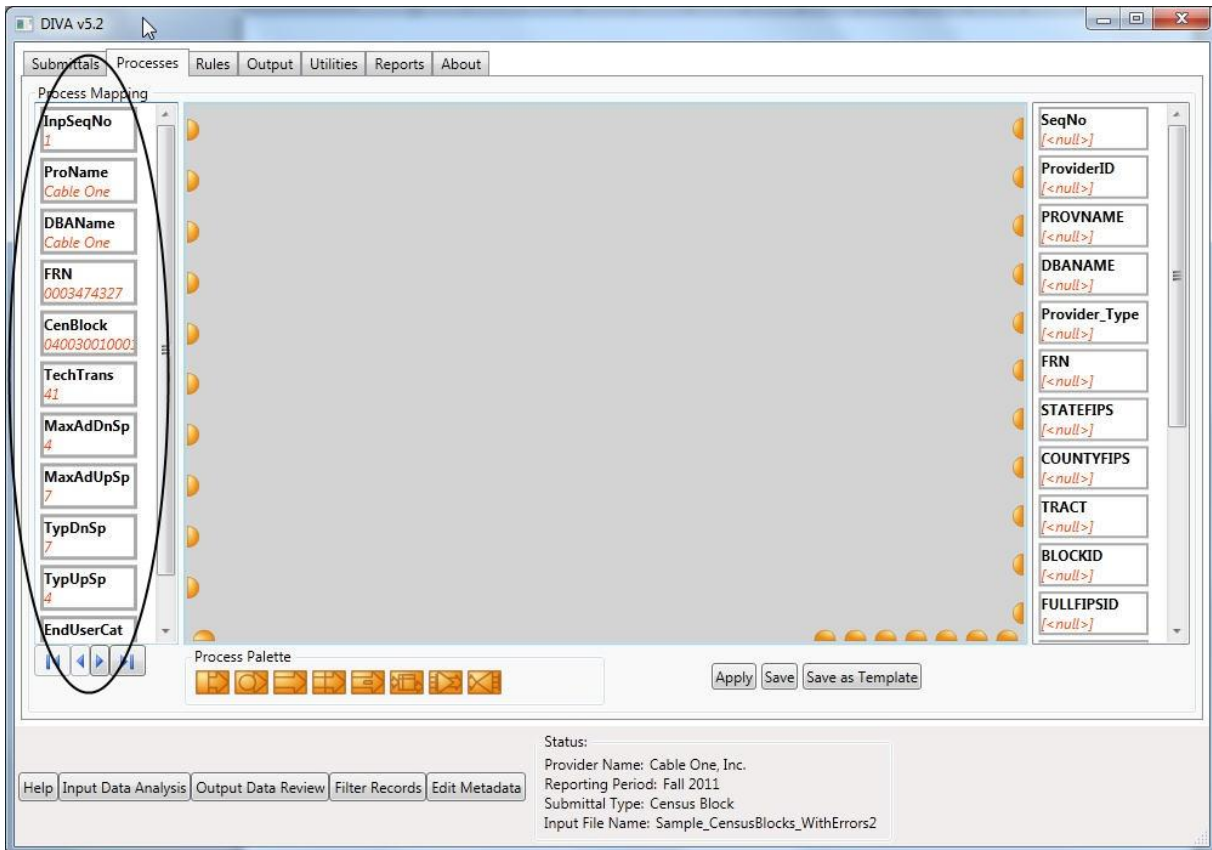
Processing Set: C1_F11_CB_8 Copy Paste

Status: Provider Name: Cable One, Inc. Reporting Period: Fall 2011 Submittal Type: Census Block Input File Name: Sample_CensusBlocks_WithErrors2

Help Input Data Analysis Output Data Review Filter Records Edit Metadata

Processes Tab Functions

Input Fields List (Processes Tab, list) – The list of Input Fields as read from the Provider Input File.



- Each “box” in the list represents one Input Field. The name assigned to that Field is shown in bold typeface and the value it has for the current Input File record is shown in red below it.
- If there was not a header row in the Input File, the Fields will be named “Col_00”, “Col_01”, etc.; if there was a header row, then the Field names will have the values obtained from that header row. In any case, the first Field will always be named “InpSeqNo”; it is the sequential “record number” that DIVA automatically creates and assigns during data loading so that the user can reliably cross-reference DIVA data with the original Provider’s data set information.

- The Input Field names in the list will reflect any changes a user may have made using the “[Input Data Analysis](#)” window. The user may also change the Input Field names here. Double clicking on the Input Field “box” will produce a dialog box where the existing name may be edited or replaced.
- If the Input Field list is longer than can be displayed on the screen, a scroll bar is shown and the half-connectors will arrange themselves along the bottom or top of the Process Canvas, depending on the position of the scroll bar. The User may use the scroll bar to the right of the Input Field List to see these hidden Input Fields, or make the DIVA window large enough to display all of them (which may not be possible if there are a large number of Input Fields).
- The orange half-connectors along the left side of the Process Mapping Canvas correspond to the Input Field with which they are adjacent; they represent data “sources” and are used to connect (link) a Process with the Input Field. A [data source](#) (or input) is a half-circle with its arc to the right.
- The color of the connector will change depending on its status with regard to the “flow” of data. If data is “flowing” out of the (associated) Input Field, due to its connection to a Process, then the connector will turn green.
- The color of the connector will also change depending on its status as an “eligible” target when the user is dragging an [Input Process Field](#). If it is “eligible” to have the Input Process Field dropped on it, then it will turn dark blue. If dropping the Input Process Field would result in a connection, then it will turn a light blue (as will the dragged Input Process Field).
- There is a “splitter” line separating the Input Fields List from the Process Mapping Canvas. Click and drag this line left, or right, to decrease, or increase, the displayed width of this list.

- A Process is created by “dragging” a Process from the “Process Palette” and dropping it onto the “Process Mapping Canvas”.
 - In the case of Decomposition Processes, a dialog box will be displayed for the user to specify which type of Decomposition Process is to be created. Note that all Decomposition Processes have the same symbology; the user can identify the Decomposition Process from its “tool tip” (see below).

- A Process can be connected by dragging an Input or Output Process Field (shown as orange half-circles with a connecting red line) to an eligible target and dropping it there. Eligible targets are shown in dark blue when the user starts the dragging operation. Both the target and the dragged Field will turn light blue when they are eligible to be connected; dropping only has an effect when both are light blue (otherwise, the user is simply dropping the Field back on the canvas). The results of dropping on an “eligible target” depends on the nature of the target:
 - If an Input Process Field is dropped on a data source (an Input Field’s connector), then the Input Process Field disappears and the connecting line attaches directly to the Input Field’s connector. In addition, the Input Field’s connector will turn dark green, the connecting line will turn light green, and the Process will turn dark green to indicate that data is now “flowing” from the data source.
 - If an Output Process Field is dropped on a data sink (an Output Field’s connector), then the Output Process Field disappears and the connecting line attaches directly to the Output Field’s connector. In addition, if the connecting line is light green (has data “flowing” into it), then the Output Field’s connector will turn dark green.
 - If an Input Process Field is dropped on an Output Process Field (or vice versa), then both will disappear and be replaced by a Process Field represented by a small circle. The respective connecting lines will now connect to this Process Field. If the incoming connecting line is light green, then the Process Field will turn dark green and the outgoing connecting line will turn light green (these color changes will “flow” down the connectors, Processes, etc., until an end is reached).
 - An Input Process Field can be dropped on eligible Process Fields (circles). If a Process Field is eligible for a drop, then it will turn dark blue when the Input Process Field is being dragged (and light blue when a drop would result in a connection). Output Process Fields can never be dropped on a Process Field (it can have only one source of input data).
 - Only Input and Output Process Fields can be “dragged and dropped” to create data connections.

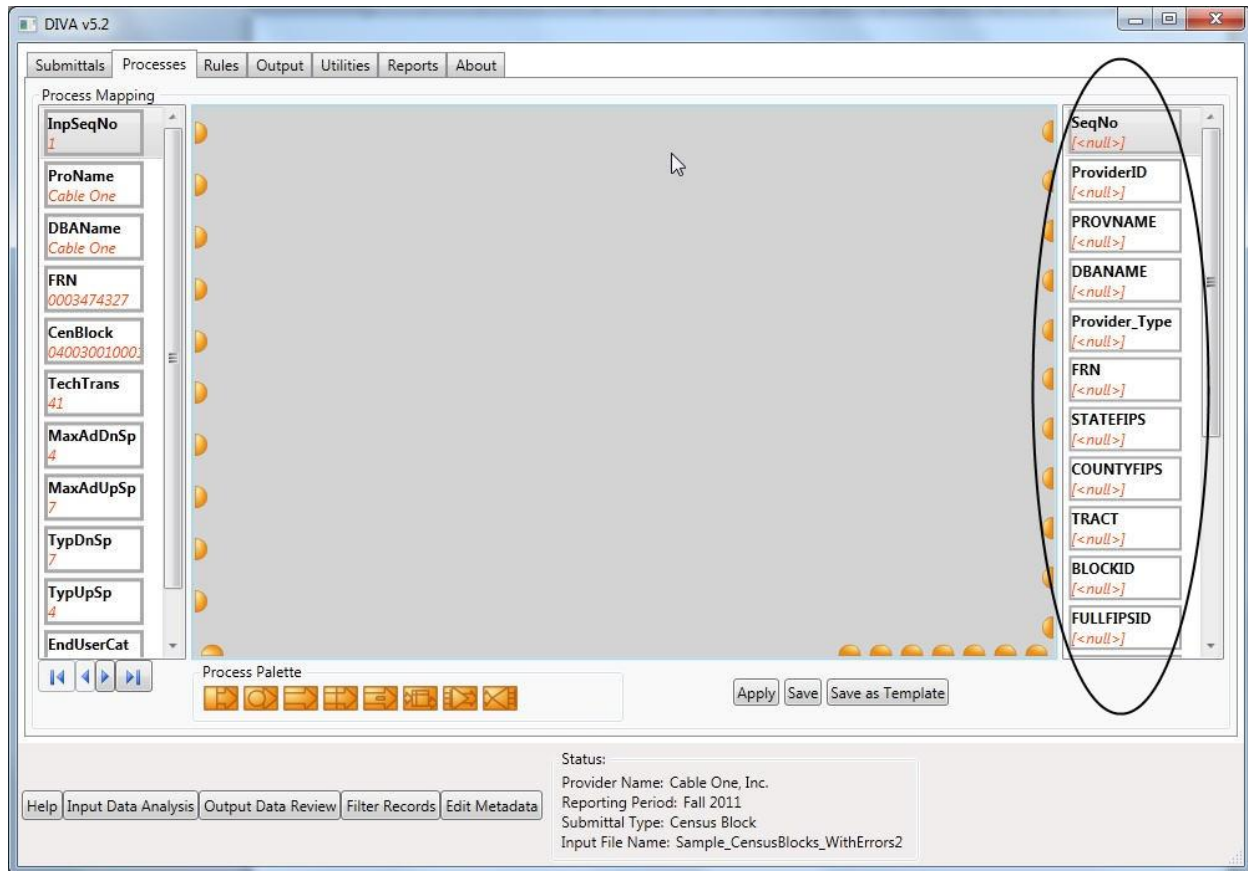
- Connected items turn from orange to dark green to give the user visual feedback that a proper connection has been made and data will flow between the green elements.

- Connectors (Input / Output Process Fields; half-circles), Process Fields (circles) and Processes (rectangles) can be selected and moved individually, or as a group by pointing and clicking for individual elements or dragging a box around multiple elements. The use of the control and shift keys as part of the selection process has the expected behavior of adding / removing items from the selection set. Elements in the selection set will have a “dotted line” shown around them.
- Processes can be connected in any allowed order to get the desired transformation of data (e.g., a string of simple decomposition processes may be required to decompose a complex string). DIVA will not permit creation of “circular” relationships; this is enforced by limiting eligible targets for dragged items.
- Processes cannot be connected directly to each other. Connections are always made to Fields (Input, Output or Process). You cannot drop a Process on anything (there are no eligible targets). You cannot drop a connector on a Process; it is never an eligible target.
- Data sources (when the half-circle arc is on the right) may have multiple connection lines running from them; that is, you may attach many connectors to a data source. Data sources can feed many Processes.
- Data sinks (when the half-circle is on the left) can have only one connection line running into it; it can receive data from only one source.
- Process Fields (circles) are the only way to connect two Processes. The Process Field is always a data sink (half-circle on the left) because it is receiving data from some Process. It is always a data source (half-circle on the right) because it is providing data to some Process. **A Process Field can only exist between two Processes;** they are only needed when Processes need to be connected to each other. They are automatically created when an Input Process Field is dropped on an Output Process Field (or vice versa).
- Processes can be disconnected by right-clicking on a Field (an Input, Process, or Output) and choosing “Detach” from the context menu. If there are multiple connections (i.e., multiple connection lines) on the Field, then all are disconnected (detached).
- Process can be deleted by selecting and then pressing the “Del” (delete) key or selecting “Delete” from the right-click context menu. Only Processes can be deleted. Multiple Processes can be deleted at one time by selecting several at once. Deleting a Process has no effect on Input Fields or Output Fields (you cannot delete these), although it may disconnect them from data “flow”.
- Processes have a varying number of Input Process Fields (half-circle on the right) and Output Process Fields (half-circle on the left). Some Processes (Constant Value and Provider Value) have no Input Process Fields in which case they serve as a data source themselves. The Composition Process allows for the addition

(and removal) of Input Process Fields, but the user can never reduce this below the minimum number of two.

- All Processes have at least one Output Process Field; Decomposition Processes have several (one for each decomposition “part”). All Processes, except Decomposition Processes, allow for the addition (and removal) of Output Process Fields, but the user can never reduce their number below one.
- Some Processes require the user to do some configuration before they have any effect. Configuration is done by double-clicking on the Process. A dialog box specific to that Process will then be displayed. Some Processes (Copy, Copy With Default) have no configuration options; double-clicking on these has no effect.
- “Tool tips” (pop-up text) are available to help users understand how to connect Processes. Just place the mouse (“hover”) over a connector or a Process to see the tool tip.

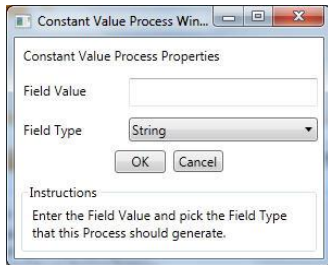
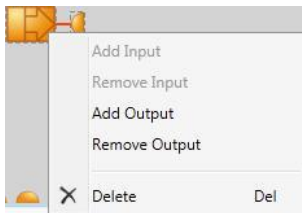
Output Fields List (Processes Tab, list) – The list of Output Fields as read from the SBDD geodatabase schema for the current Reporting Period and Submittal Type (Period-Type).



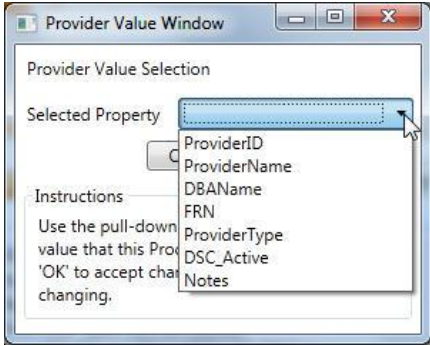
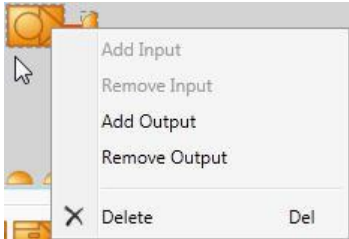
- Each “box” in the list represents one Output Field. The name assigned to that Field is shown in bold typeface and the value it has for the current Input File record, as transformed by the Processes, is shown in red below it. In the illustration above, no Processes have been defined; therefore all values in the Output Fields are null.
- The user can see the effect of the assigned Processes on data transformation by examining these values. The values can be changed by adjusting the Processes as needed.
- The name assigned to each Output Field in this list is not editable; they come from the SBDD geodatabase and are the NTIA-approved field names (or, in the case of DIVA-related Output Fields, from the standard field names assigned by the DIVA design team).

- The orange half-connectors along the side of the Process Mapping Canvas correspond to the Output Field with which they are adjacent; they represent data “sinks” and are used to connect (link) a Process with the Output Field. A data sink (or output) is a half-circle with its arc to the left. Unlike Input Fields, Output Fields cannot have more than one connection to Processes. It is likely that one to several Output Fields will have no connections as the data is not present in the Input File and cannot be filled in using one of the Processes that generates values (i.e., Provider Value Process).
- If the Output Field list is longer than can be displayed on the screen (i.e., a scroll bar is shown), then these orange half-connectors will arrange themselves along the bottom (or top, depending on the position of the scroll bar). The User may use the scroll bar to the right of the Output Field List to see these hidden Output Fields, or make the DIVA window large enough to display them all of them (which may not be possible if there are enough Output Fields).
- The color of the connector will change depending on its status with regard to the “flow” of data. If data is “flowing” into the (associated) Output Field, due to its connection to a Process, then the connector will turn green.
- The color of the connector will also change depending on its status as an “eligible” target when the user is dragging an Output Process Field. If it is “eligible” to have the Output Process Field dropped on it, then it will turn dark blue. If dropping the Output Process Field would result in a connection, then it will turn a light blue (as will the dragged Output Process Field).
- The area below the Output Fields List is used to display any “Row” Processes that may exist.
- There is a “splitter” line separating the Output Fields List from the Process Mapping Canvas. Click and drag this line left, or right, to increase, or decrease, the displayed width of this list.


Constant Value Process (Processes Tab, Process Palette) – Used to create a user-defined constant value. This Process is useful when there is no Input Field to provide data for a required Output Field. A good example of this is the numeric code for End User required in recent versions of the SBDD geodatabase. It is rare that a Provider submits this value and the Constant Value Process can be used to put in the appropriate value (e.g. a “5” representing “Other”), when a user does not have any other information.

Action	Result
Drag and drop the Constant Value Process onto the Process Canvas from the Process Palette	The control appears on the canvas and is ready to be configured. By default, it will have one Output Process Field; others may be added. It will never have an Input Process Field as it serves as its own data source.
Double-click the Constant Value Process	<p>A configuration dialog box is opened. The user types in a value and specifies the value’s type (e.g. string).</p> 
Right-click the Constant Value Process on the Process Mapping Canvas	<p>May add or remove outputs. The constant value can be sent to one or more Output Fields or Processes.</p> 

Provider Value Process (Processes Tab, Process Palette) – Used to create a constant value based on a specific Provider attribute. This Process is useful to ensure that values for Provider ID, Provider Name, DBA Name, FRN and Provider Type (Output Fields required by most Period-Types) are standardized from a table imported into DIVA

Action	Result
Drag and drop the Provider Value Process onto the Process Mapping Canvas from the Process Palette	The control appears on the canvas and is ready to be configured. By default, it will have one Output Process Field; others may be added. The Process box will be orange, until the user configures its properties.
Double-click the Provider Value Process	<p>A configuration dialog box opens which allows the user to select the field in the Provider Information Table for the active Provider that will populate the Output Field.</p> 
Right-click the Provider Value Process on the Process Mapping Canvas	<p>May add or remove outputs. The constant value can be sent to one or more Output Fields or Processes. The number can never be reduced below one. This means that a single Provider value can be directed to one or more Output Fields and/or Processes.</p> 

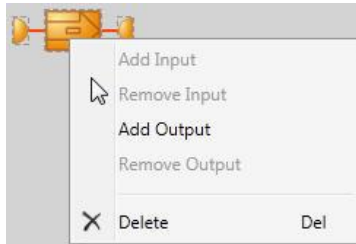
Copy Process (Processes Tab, Process Palette) – this Process populates an Output Field or Process by copying the value of the Input Field or Process.

Action	Result
Drag the Copy Process onto the Process Mapping Canvas	The control appears on the canvas and is ready to be configured. It will be orange, until the user connects its input side to an Input Field or another Process and its output side to an Output Field or another Process.
Double-click the Copy Process	Note that there is no configuration dialog for this Process. Just drag, drop and connect. No transformative process is applied other than to cast the Provider value into the field type of the Output Field
Right-click the Copy Process on the Process Mapping Canvas	May add or delete outputs. 

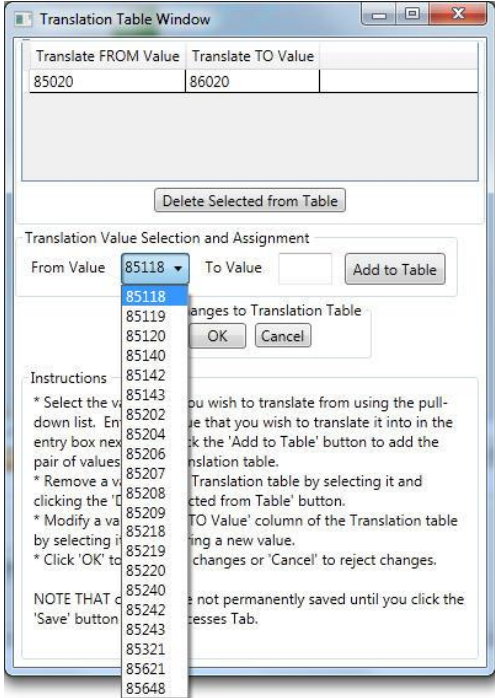
Copy with Default Process (Processes Tab, Process Palette) – this Process populates an Output Field or Process with a default value as defined in the active SBDD database, where the input values are either null or blank.

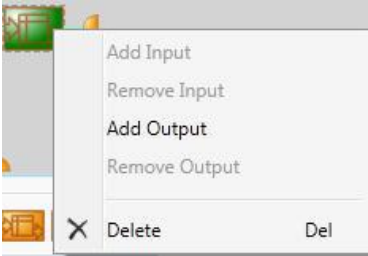
Action	Result
Drag the Copy with Default Process onto the Process Mapping Canvas	The control appears on the canvas and is ready to be configured
Double-click the Copy with Default Process	Note that there is no configuration dialog for this Process. The value is determined by the default value contained in the SBDD geodatabase.
Right-click the Copy with Default Process on the Process Mapping Canvas	Note that no configuration options are available for this Process. The User cannot specify multiple input or output fields.

Make Negative Process (Processes Tab, Process Palette) – this Process inverses the sign of the value coming from a Input Field or Process.


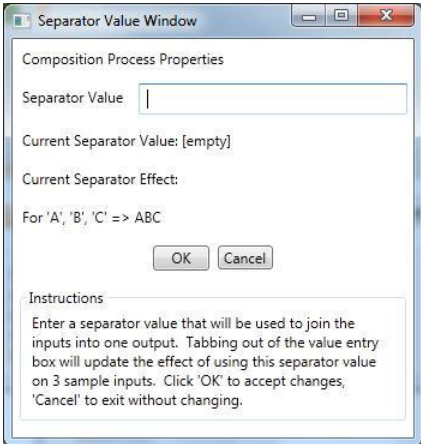
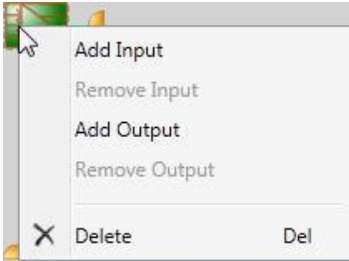
Action	Result
Drag the Make Negative Process onto the Process Mapping Canvas	The control appears on the canvas and is ready to be configured.
Double-click the Make Negative Process	Note that there is no configuration dialog for this Process.
Right-click the Make Negative Process on the Process Mapping Canvas	<p>Add or remove an output connection.</p> 

Translation Process (Processes Tab, Process Palette) – this Process is used to translate one or more values from an Input Field or Process into another value in the Output Field.

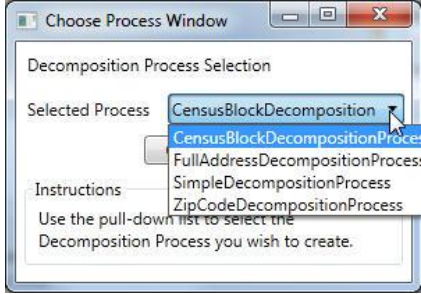
Action	Result
Drag the Translation Process onto the Process Mapping Canvas	The control appears on the canvas and is ready to be configured.
Double-click the Translation Process	<p>A configuration dialog window opens which allows the user to map values from the data source (either an Input Field or a Process Field) to new values in the data sink (either an Output Field or a Process Field).</p> <p>In the illustration below the value of “85020” is mapped to “86020” and is shown in the top part of the “Translate Table Window”.</p>  <p>The unique list of values shown in the “From Value” pull-down list comes from the data source (the Input Field or Process Field) to which the Translate Process is attached.</p> <p>Input values are mapped by typing a new value into the “To Value” text box and then clicking the “Add to Table” button. The user must click the “OK” button in the “Save</p>

	<p>Changes to Translation Table” area or click “Cancel” to exit without saving changes since the last Save.</p> <p>Values which are <u>not</u> translated are simply copied from the data source to data sink.</p>
<p>Right-click the Translation Process on the Process Mapping Canvas</p>	<p>Add or remove an output connection.</p> 

Composition Process (Processes Tab, Process Palette) – this Process is used to combine values from two or more Input Fields or Processes into a single Output Field.

Action	Result
Drag the Composition Process onto the Process Mapping Canvas	<p>The control appears on the canvas and is ready to be configured. By default, two input connectors are provided, but more can be added (see below)</p> 
Double-click the Composition Process	<p>A configuration dialog opens which allows the user to specify a separator value to be placed between the appended Input Field values.</p> 
Right-click the Composition Process	<p>Add or remove both input and output connections.</p> 

Decomposition Process (Processes Tab, Process Palette) – this Process is used to parse values from a single Input Field or Process into two or more Output Fields or Processes.

Action	Result
<p>Drag the Decomposition Process onto the Process Mapping Canvas and choose the decomposition type.</p>	<p>A dialog appears on the canvas and asks the user to specify the type of Decomposition Process to use:</p>  <p>Census Block – Decomposes a full census block ID into its four component parts (State, County, Tract and Block). There are one input and four output connectors provided. This is fixed and cannot be changed from the context menu. Tool Tips on the output connectors guide user as to which of the four parts of a full Census ID it should be connected.</p> <p>Full Address –The full address decomposition process is not active at the current release.</p> <p>Simple – Parses a string into two parts based on a user-defined symbol (e.g. a space). There are one input and two output connectors provided. This is fixed and cannot be changed from the context menu. Multiple parsing can be accomplished by chaining multiple simple decomposition Processes together.</p> <p>Zip Code – Parses a nine-digit zip code string into a five-part and four-part string. There are one input and two output connectors provided. This is fixed and cannot be changed from the context menu.</p>

<p>Double-click the Decomposition Process</p>	<p>Simple – Only the Simple decomposition type has a configurable dialog. Type the character on which the string will be parsed into the “Symbol Value” text box. The current value is displayed immediately below.</p> <div data-bbox="748 333 1036 632" data-label="Image"> </div>
<p>Right-click the Decomposition Process</p>	<p>None of the decomposition process types are configurable from the context menu</p>

Saving a Template Processing Set (Processes Tab) – the “Save as Template” button saves the active Processing Set to a Template Processing set, overwriting any pre-existing set for the current Period-Type. No GUI is involved, but there are some things to be aware of regarding these templates.

Doesn't make sense to copy between Reporting Periods due to potential changes in Output Fields. That is why this option only works within a single Reporting Period

Users can choose which set of Processes to apply within the same Reporting Period. This can be done in one of two ways. A Template Processing Set can be applied when a new Input File is created or it can be copied and pasted from an existing Processing Set to another Processing Set at any time. The workflow for each is listed below:

Applying to a New Submittal

[HW: write this up and include screen dump(s)]

Copying and Pasting a Template Processing Set

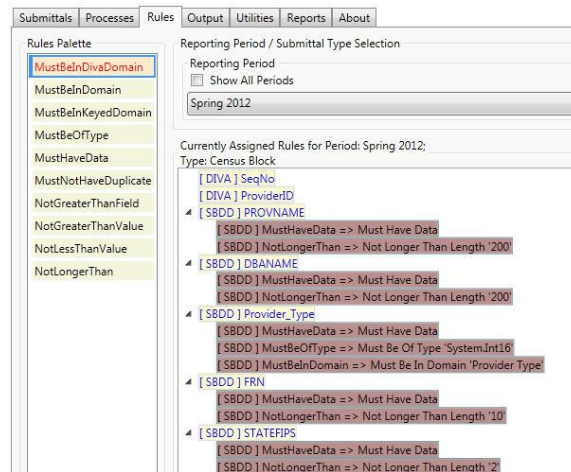
[HW: write this up and include screen dump(s)]

Rules Tab Functions

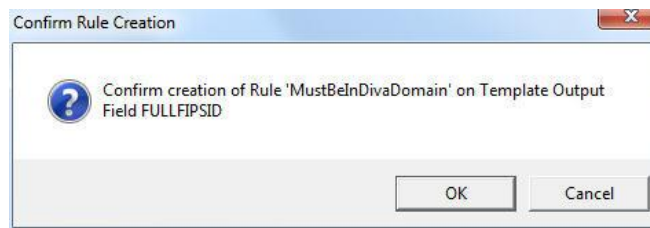
Using the Rules Palette Area (Rules Tab, “Rules Palette” control) – A list of Rule types that can be applied to the Template Output Fields by dragging and dropping them onto the Template Output Fields shown in the “Rule Summary Area”.



- As the user clicks on each rule type in the Rules Palette, the Output Fields that can accept that rule type are highlighted in blue.



- When the Rule is dropped on a valid Output Field, the user is prompted whether this is what should be done. Click “OK” or “Cancel”.

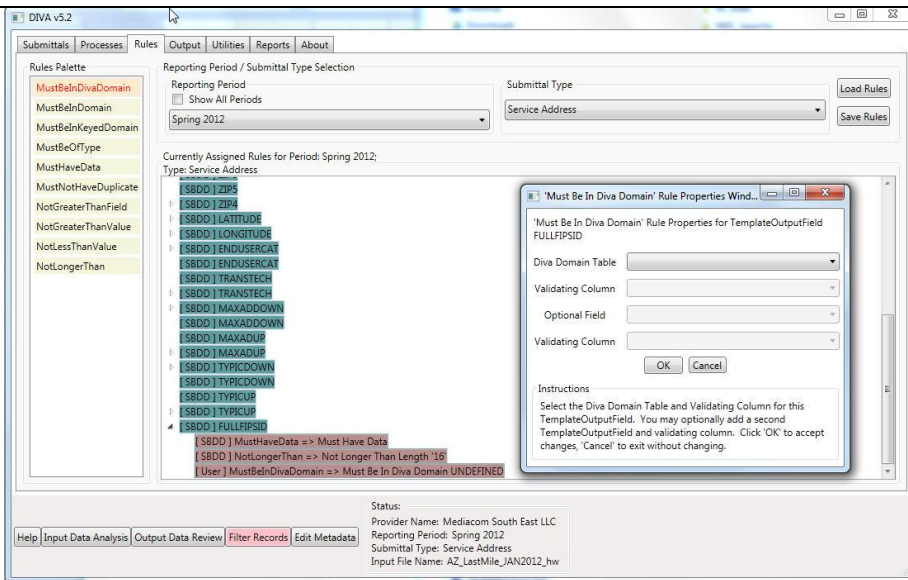


- Rules can be removed from an output field by right-clicking and selecting “Delete”. This includes Rules imported from the SBDD database. Be careful, there is no “Undo”.
- Following is a summary of each rule type. More detailed information on how to configure each rule can be found by clicking on the hyperlinks.
 - [Must Be In Diva Domain](#) – The value in the Output Field must be in a [DIVA domain](#).
 - [Must Be In Domain](#) – The value in the Output Field must be in an SBDD-defined domain.
 - [Must Be In Keyed Domain](#) – The value in the Output Field must be in an SBDD-defined [keyed domain](#).
 - [Must Be Of Type](#) – The value in the Output Field must be of a certain type. This rule type applies only to numeric fields.
 - [Must Have Data](#) – The value in the Output Field must not be empty strings or null values. There must be a value present.
 - [Must Not Have Duplicate](#) – The value in the Output Field must not result in a duplicate row.
 - [Not Greater Than Field](#) – The value in the Output Field must not be greater than the value of those in another user-specified Output Field.
 - [Not Greater Than Value](#) – The value in the Output Field must not be greater than a user-specified value.
 - [Not Longer Than](#) – The value in the Output Field must not exceed a user-defined number of characters. This rule type applies only to character fields, not numeric fields.

Using the Currently Assigned Rules Area (Rules Tab, “Rules Palette” control) – A display of the Rules assigned to each Template Output Field of a Submittal, as well as a place to drag and drop new Rules onto Template Output Fields, configure existing Rules, or remove Rules. As discussed in the introduction to the Rules tab, this area is used to assign or remove rules from the Template Output Fields and therefore any change you make here now apply to any Submittal that uses the same Template Output Field. Therefore, if a user makes changes here, they should re-run previous Submittals again to see how the new rule applies to them.

- At a minimum, the Rules for the Output Fields of a Submittal’s Input File will be those imported from the SBDD geodatabase. Remember that these will be stored on the Template Output Fields, so they will automatically apply to all Output Fields for that Period-Type. Additional Rules may be assigned by dragging a rule type from the “Rules Palette” and dropping it on a Template Output Field.
- There are some restrictions on which Rule types may be placed on which Template Output Fields. For example, a Rule type that only works with character (string) field cannot be dropped on a numeric field. If a Rule from the palette is valid for a given Template Output Field, the cursor will change to a small box with an arrow as the user drags it over the Output Field Name. Otherwise, the cursor will display as a circle with a bar through it. In addition, at the start of the “drag” operation, all eligible target Template Output Fields will change color to indicate that they will accept the Rule “drop”.
- There are also some restrictions, depending on the Rule, as to the number of the Rules that can be assigned to a Template Output Field. For example, the “Must Have Data” Rule can only be assigned to each Template Output Field one time. If a Rule has been dragged from the palette and there are no eligible targets, a dialog box will inform the user of that fact and the Rule will be made inaccessible.
- By default the rules are collapsed and all that shows in the “Currently Assigned Rules” area are the Template Output Fields for the Submittal. Rules can be viewed for a given Template Output Field by double-clicking the field name. All the fields can be expanded to show their rules by double clicking the top-most Template Output Field in the list. Double-clicking again will collapse the expanded nodes and open the collapsed nodes.
- Rules may be deleted from a Template Output Field by right-clicking on the rule and selecting the “Delete” option, or by clicking on the Rule (selecting it) and then pressing the “Del” (delete) key.
- The MustNotHaveDuplicate Rule can only be applied once. After it is applied, that Rule choice will be greyed out in the palette.
- If any Rule no longer has eligible targets, that Rule will be de-activated (greyed out) in the palette. To re-activate it, you must delete at least one occurrence of that Rule.

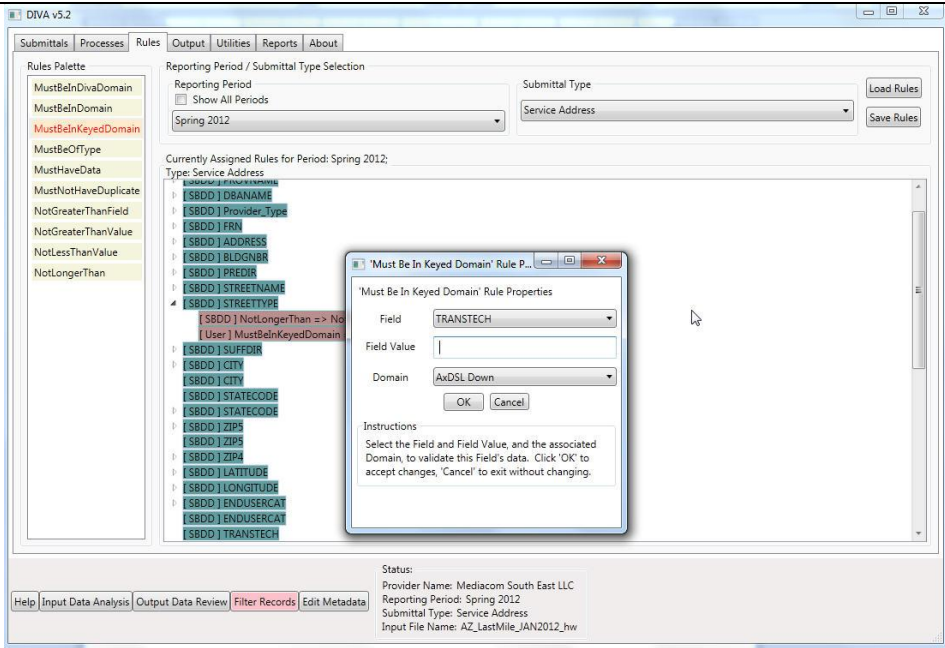
MustBeInDIVADomain Rule (Rules Tab, “Rules Palette” control) – allows users to apply their own domains to Output Field values. These domains are called “DIVA Domains” to distinguish them from the domains imported from the SBDD geodatabase.

Action	Result
Import one or more domains from the “Utilities” tab.	A domain table must be imported into DIVA before it will be available to be used in a Diva Domain rule. This procedure is documented in the “Manage DIVA Domains” section of this User Manual.
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	 <p>The screenshot shows the DIVA v5.2 application window. On the left is the 'Rules Palette' with various rules listed. The 'Must Be In Diva Domain' rule is highlighted. In the center, a list of 'Currently Assigned Rules for Period: Spring 2012' is shown, with the 'Must Be In Diva Domain' rule selected. On the right, the 'Must Be In Diva Domain' Rule Properties window is open, showing fields for 'Diva Domain Table', 'Validating Column', and 'Optional Field'. The 'Diva Domain Table' is set to 'FULLFIPSID'. The 'Validating Column' is set to 'FULLFIPSID'. The 'Optional Field' is set to 'FULLFIPSID'. The 'Rule Properties' window also contains instructions: 'Select the Diva Domain Table and Validating Column for this TemplateOutputField. You may optionally add a second TemplateOutputField and validating column. Click 'OK' to accept changes, 'Cancel' to exit without changing.'</p> <ul style="list-style-type: none"> • When the Rule is first placed under an Output Field, the text “UNDEFINED” will appear next to it. This means the user must configure the rule before it can be applied. • Select the imported domain from the “DIVA Domain Table” pull-down. • Select the field in the imported domain table that has the domain value from the “Validating Column” pull-down. • Optionally select an additional Template Output Field and a validating column from this DIVA Domain table; this allows the user to check the validity of two Template Output Fields against the same table (e.g. check both Census Block ID and Census Block area values simultaneously). • Click “OK” to apply the domain, or “Cancel” to exit without applying the domain.

MustBeInDomain Rule (Rules Tab, “Rules Palette” control) – allows users to apply an SBDD-generated domain to Output Field values. These domains are called “Domains” to distinguish them from the “DIVA Domains” imported by users from local source tables. This rule type allows users to apply any existing SBDD Domain to any compatible Output Field, whether it makes sense or not. It is likely this rule type will not be applied very often, if at all, but it is another tool DIVA users have to help validate Provider data.

Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	<div data-bbox="386 709 1393 1396"> </div> <ul style="list-style-type: none"> When the Rule is first placed under an Output Field, the text “UNDEFINED” will appear next to it. This means the user must configure the rule before it can be applied. Select the SBDD domain from the “Domain” pull-down. Click “OK” to apply the domain, or “Cancel” to exit without applying the domain.

MustBeInKeyedDomain Rule (Rules Tab, “Rules Palette” control) – allows users to apply an SBDD-generated [Keyed Domain](#) to Output Field values. These domains are called “Domains” to distinguish them from the “DIVA Domains” imported by users from local source tables. This rule type allows users to apply any existing SBDD Keyed Domain to any compatible Output Field, whether it makes sense or not. It is likely this rule type will not be applied very often, if at all, but it is another tool DIVA users have to help validate Provider data.

Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	 <ul style="list-style-type: none"> • When the Rule is first placed under an Output Field, the text “UNDEFINED” will appear next to it. This means the user must configure the rule before it can be applied. • Double-click the Rule to open the configuration dialog • Select the Output Field containing the key value which will be used to look up the appropriate values in the Domain • Type in a key value found in the Output Field which will be used to look up the appropriate values in the Doman • Select the appropriate SBDD Domain from the “Domain” pull-down. • Click “OK” to apply the domain, or “Cancel” to exit without applying the domain.


MustBeOfType Rule (Rules Tab, “Rules Palette” control) – allows users to check the Output Field value type. This rule is only valid for numeric Output Fields.

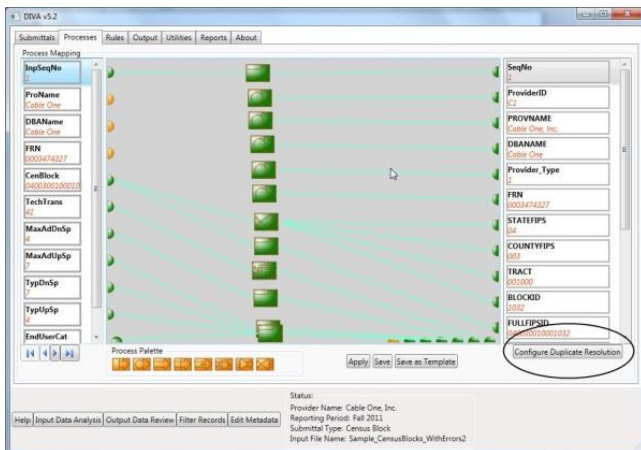
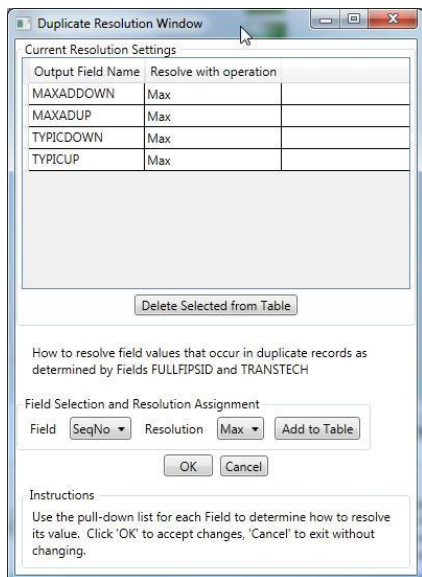
Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	There are no configuration options for this rule.

MustHaveData Rule (Rules Tab, “Rules Palette” control) – allows users to check the Output Field for empty or null value. If found, the Rule throws an error.


Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	There are no configuration options for this rule.

MustNotHaveDuplicate Rule (Rules Tab, “Rules Palette” control) – allows users to check Output Fields for values that would result in a duplicate row. This is a very different type of Rule than all the others. It operates across rows and not on a column or column(s), exclusively.

Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	 <ul style="list-style-type: none"> • The target field on which the “MustNotHaveDuplicates” Rule was dropped was the census block ID (FULLFIPSID), so that part of the configuration is done. • This dialog is asking for a second field, the one that contains the value that should not repeat. In this case the user specified the technology type (TRANSTECH) from the pull down meaning that for any unique FULLFIPSID, there cannot be two identical TRANSTECH values.

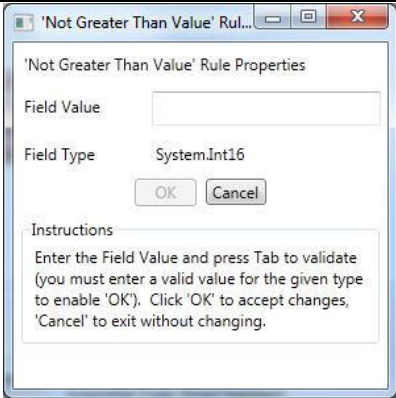
Action	Result
	<div></div> <ul style="list-style-type: none">Clicking on the “Configure Duplicate Resolution” button opens the dialog below. <div></div> <ul style="list-style-type: none">Using the pull-down lists at the bottom of the dialog, specify how each of the field values for duplicate records should be handled.In the example above, each of the four speed fields is told to take the maximum value found in any of the duplicate records and to use that for the final value in the single combined record that will be created.

NotGreaterThanField Rule (Rules Tab, “Rules Palette” control) – the value in the Output Field must not be greater than the value in another user-specified Output Field.


Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	 <ul style="list-style-type: none"> • Select the “not greater than field” from the pull-down list. The values in this field must be less than the values in the rule field or the record is flagged as an error.

NotGreaterThanValue Rule (Rules Tab, “Rules Palette” control) -- the value in the Output Field must not be greater than a user-specified value.

Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.

Action	Result
Double-click the Rule under the Output Field to Configure	 <ul style="list-style-type: none"> Type in the “Not Greater Than Value” and press tab key to validate. Select “OK” to accept or “Cancel” to exit without setting/changing the value.

NotLongerThan Rule (Rules Tab, “Rules Palette” control) – the value in the Output Field must not exceed a user-defined number of characters. This rule can only be applied to character (string) fields.

Action	Result
Drag and drop the Rule onto the target field	When the “drag” of a Rule is initiated from the Rules Palette, the acceptable Template Output Field targets will change color. As noted, DIVA has certain restrictions on the number and types of Rules that can be assigned to various Template Output Fields.
Double-click the Rule under the Output Field to Configure	 <ul style="list-style-type: none"> Type in the “Not Longer Than” value and click “OK” to accept or “Cancel” to exit without setting/changing the value.

Output Tab Functions

Rule Violations and Counts area (Output Tab) - Displays the results of applying Rules to the Output Fields of a Submittal.

Rule Violations and Counts

- [SBDD] TRACT
 - [SBDD] MustHaveData 11
 - [SBDD] NotLongerThan 0
- [SBDD] BLOCKID
 - [SBDD] MustHaveData 0
 - [SBDD] NotLongerThan 0
- [SBDD] FULLFIPSID
 - [SBDD] MustHaveData 0
 - [SBDD] NotLongerThan 0
- [User] MustNotHaveDuplicate 23
- [SBDD] TRANSTECH
 - [SBDD] MustHaveData 0
 - [SBDD] MustBeOfType 0
- [SBDD] MAXADDOWN
 - [SBDD] NotLongerThan 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0
 - [SBDD] MustBeInKeyedDomain 0

Failed Record View for FULLFIPSID => Must Not Have Duplicate Records with Field TRANSTECH

InpSeqNo	FULLFIPSID	TRANSTECH	MAXADDOWN	MAXADUP	TYPICDOWN	TYPICUP
304	040030011003032	40	7	4	6	2
305	040030011003032	40	6	4	7	4
254	040030011003047	40	7	4	7	3
255	040030011003047	40	7	3	7	4
256	040030011003053	40	5	4	7	5
257	040030011003053	40	6	2	7	7
258	040030011003053	40	7	4	7	
259	040030011003053	40	5	3	7	4
260	040030011003059	40	7		7	4
261	040030011003059	40	4	5	7	4
262	040030011003059	40	5	4	7	3
263	040030011003059	40	6	2	7	4
264	040030011003059	40	7	3	7	
301	040030012002012	40	7	4	5	4
302	040030012002012	40	6	3	6	2
303	040030012002012	40	7	4	7	3
306	040030012002019	40	7	3	7	3
307	040030012002019	40	6	3	6	3
308	040030012002019	40	5	4	4	4

Duplicate rows as output.

InpSeqNo	FULLFIPSID	TRANSTECH	MAXADDOWN	MAXADUP	TYPICDOWN	TYPICUP
304	040030011003032	40	8	8	9	4
254	040030011003047	40	7	4	7	4
256	040030011003053	40	7	4	7	

Status:
 Provider Name: Cable One, Inc.
 Reporting Period: Fall 2011
 Submittal Type: Census Block
 Input File Name: Sample_CensusBlocks_WithErrors2

- This is a list of Output Fields with each Rule that has been assigned to the associated Template Output Field shown underneath it. The number of Rule violations (errors) is shown to the immediate right of each Rule.
- By default, the Rules under each Output Field are not visible, having been collapsed under each Output Field. The user can toggle the visibility of rules by double-clicking on the Output Field. The visibility of all rules can be toggled by double-clicking on the top-most Output Field in the list.
- Each Rule has either “[SBDD]” or “[User]” listed before the Rule name. This reflects whether the Rule was added as part of the SBDD XML Importoperation or subsequently by the user.

- Placing and holding (“hovering”) the mouse over any Rule will display the more informative Rule description as a tool-tip.
- A Rule violation count of “-2” indicates that the Rule is not configured and therefore cannot be evaluated. Rules are configured on the Rules Tab.
- A Rule violation count of “-1” indicates that there is no data in the Output Field to be evaluated. Either the Output Field is not connected to a data source on the Processes tab or there is no data in the connected data source.
- More information about errors can be displayed by double-clicking on a Rule that is showing one or more errors. The “Failed Record View” will be populated with the rows that were flagged with errors in the Output Field.
- Users work iteratively, adding/deleting/editing Processes, Rules and Filters until zero errors are achieved or the errors that remain are acceptable. For example, there may be no address information and no way to get it, so those Output Fields will always show errors in the MustHaveData rule.

Failed Record View area (Output Tab) – Displays the rows which have an error in the Output Field for a Rule which has been double-clicked on by the user in the “Rules and Violations Count” area.

The screenshot shows the DIVA v5.2 application window. The 'Output' tab is selected. On the left, the 'Rule Violations and Counts' panel shows a tree view of rules. The 'SBDD] TRANSTECH' rule is selected, showing a count of 23 violations. The main panel displays a 'Failed Record View for FULLFIPSID => Must Not Have Duplicate Records with Field TRANSTECH'. This panel contains a table with columns: InpSeqNo, FULLFIPSID, TRANSTECH, MAXADDOWN, MAXADUP, TYPICDOWN, and TYPICUP. The table lists 23 records. The first record (InpSeqNo 305) is highlighted in cyan, indicating it is the first non-error record of a duplicate set. The subsequent 22 records are highlighted in red, indicating they contain errors. Below the main table, there is a section titled 'Duplicate rows as output' which shows a single record for each duplicate set, with values adjusted according to the 'Resolve Duplicate' process settings. The bottom status panel shows the provider name 'Cable One, Inc.', reporting period 'Fall 2011', and input file name 'Sample_CensusBlocks_WithErrors2'.

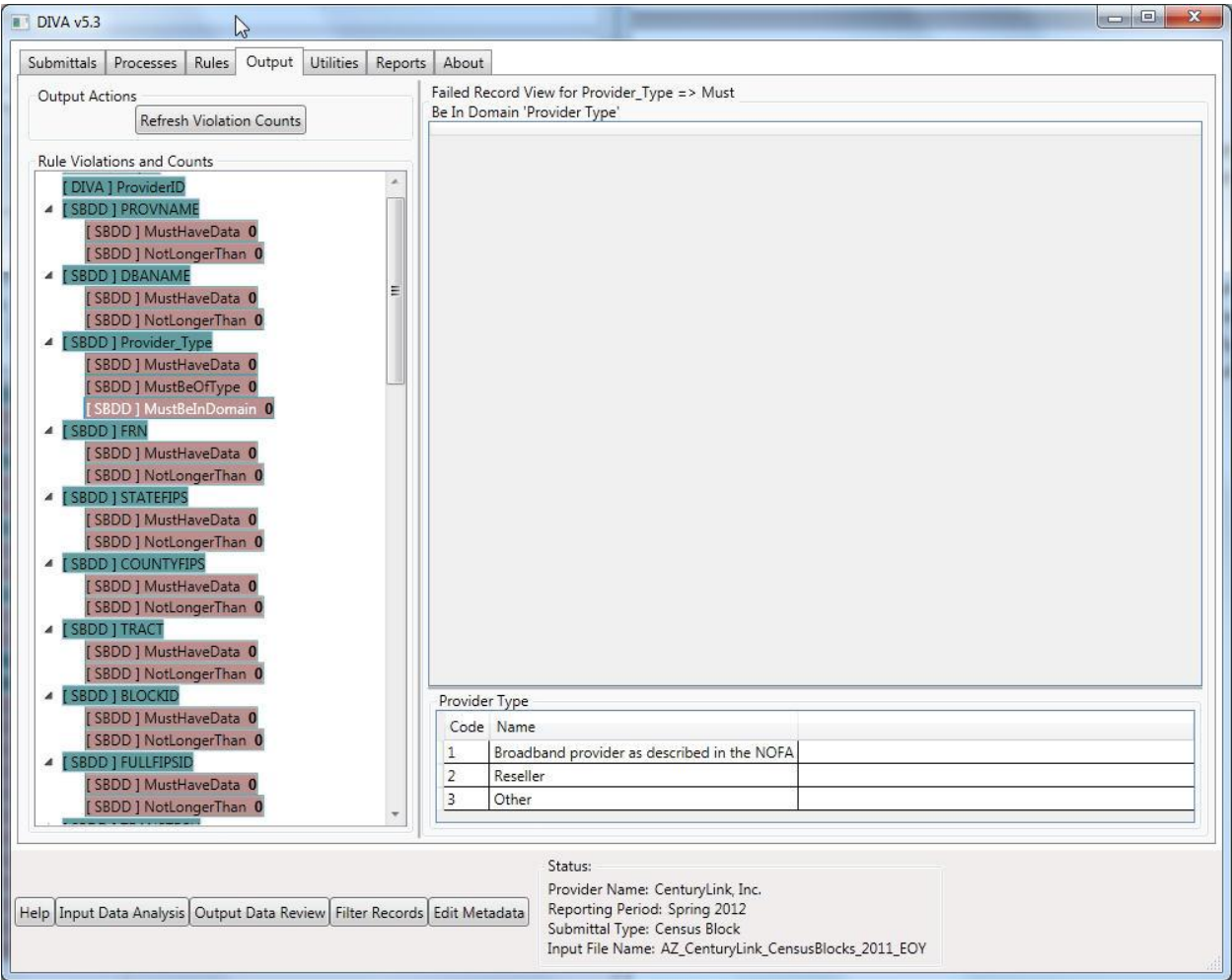
InpSeqNo	FULLFIPSID	TRANSTECH	MAXADDOWN	MAXADUP	TYPICDOWN	TYPICUP
305	040030011003032	40	7	4	6	2
254	040030011003047	40	7	4	7	3
255	040030011003047	40	7	3	7	4
256	040030011003053	40	5	4	7	5
257	040030011003053	40	6	2	7	7
258	040030011003053	40	7	4	7	
259	040030011003053	40	5	3	7	4
260	040030011003059	40	7		7	4
261	040030011003059	40	4	5	7	4
262	040030011003059	40	5	4	7	3
263	040030011003059	40	6	2	7	4
264	040030011003059	40	7	3	7	
301	040030012002012	40	7	4	5	4
302	040030012002012	40	6	3	6	2
303	040030012002012	40	7	4	7	3
306	040030012002019	40	7	3	7	3
307	040030012002019	40	6	3	6	3
308	040030012002019	40	5	4	4	4

InpSeqNo	FULLFIPSID	TRANSTECH	MAXADDOWN	MAXADUP	TYPICDOWN	TYPICUP
304	040030011003032	40	8	8	9	4
254	040030011003047	40	7	4	7	4
258	040030011003053	40	7	4	7	

Status:
 Provider Name: Cable One, Inc.
 Reporting Period: Fall 2011
 Submittal Type: Census Block
 Input File Name: Sample_CensusBlocks_WithErrors2

- This display will change somewhat depending on the error type being displayed.
 - In the illustration above, “MustNotHaveDuplicates” errors are being displayed with records that have errors being shown in red. The first non-error record of each duplicate set (if there is one) will be shown in cyan (light blue); this record will be exported (however, not necessarily with the values shown in this panel). Subsequent non-error records of each duplicate set of records will be shown in darker blue; these will not be exported.
 - The rows that will be exported from DIVA are shown in the Extra Data panel (labeled, in this case, with “Duplicate Rows as output”). There will be one record here for every record in the above panel that is in cyan. However, their values have been adjusted in accordance with the settings of the “Resolve Duplicate” Process (on the Processes Tab, below the Output Fields List). In this example, the “Resolve Duplicate” Process is set to resolve values by selecting the maximum value of each of the four speed fields.

Domain Table Display area (Output Tab) – Displays the values in the domain table for a Rule which has been double-clicked on by the user in the “Rules and Violations Count” area.



In the illustration above, “MustBeInDomain” rule has been clicked with the result that (a) the rule changes to white font and (b) the domain table for that rule is being displayed below the “Failed Record View” area. The “Failed Record View” is blank in this case, because no records violated the domain Rule

Utilities Tab Functions

Import SBDD XML File (Utilities tab, “Import SBDD XML File” control)

The “Import SBDD XML File” control allows the user to import a file (text, dbf or Excel) containing SBDD database schema and domain values into a DIVA table.

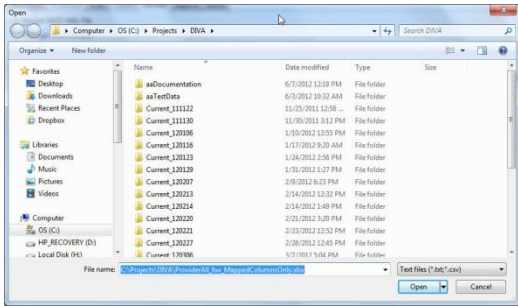
THIS STEP MUST BE PERFORMED FOR A GIVEN REPORTING PERIOD BEFORE ANY OTHER PROCESSING CAN TAKE PLACE.

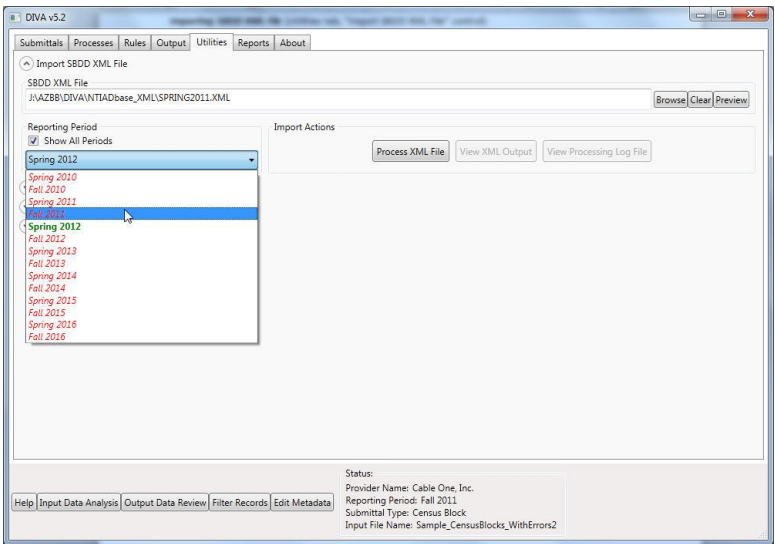
This step creates the definitions for most Template Output Fields (DIVA has some extended fields for various SBDD feature classes which are added by DIVA automatically) and imports the Rules for them by processing the SBDD domains.

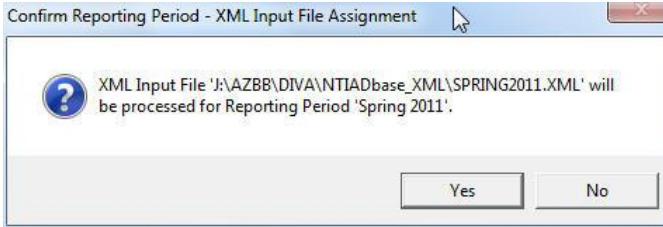
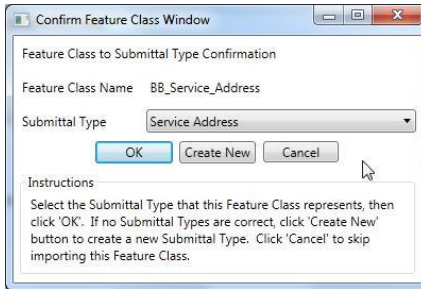
A pre-cursor step for this function is to use ArcCatalog to export a workspace XML document from an SBDD geodatabase.

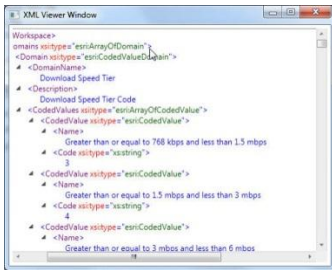
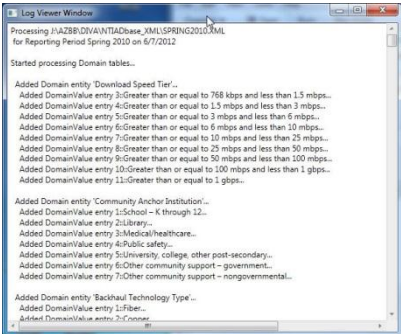
Click on “Import SBDD XML File” section of Utilities Tab and the “Import SBDD XML File” area expands:

Key	Control	Description
1	"SBDD XML File" text box	Displays the path and file name to the SBDD XML file.

Key	Control	Description
2	Browse button	<p>Click on the “Browse” button and the browse dialog box opens:</p>  <p>Choose the file type filter at the bottom of the Browse dialog to match the file type (extension) of your Provider information file.</p> <p>Select the source file and click “Open” on the file dialog box. The browse dialog goes away and you will be back to the Utilities tab with the pathname and file to your source file now showing in the “SBDD XML File” text box.</p>
3	Clear button	<p>Clears the value in the “SBDD XML File” text box. Does not delete the XML file or unload a previously loaded XML file. It just clears the text box.</p>
4	Preview Button	<p>Opens the DIVA “Preview Window” which has been previously discussed.</p>

Key	Control	Description
5	Reporting Period area	<p>Click on the “Reporting Period” pull-down menu and the list of Reporting Periods is displayed. The currently active Reporting Period is shown in green while the rest of them are in red font.</p>  <p>Select the appropriate Reporting Period from the pull down list for the selected XML Import File. The active Reporting Period is shown in green font.</p> <p>If an XML file for a given Reporting Period has already been imported, the “Process XML File” button will be inactive.</p> <p>The selection of the Reporting Period and the source XML file can be done in any order.</p> <p>Note: the pull down list can be managed (to a certain extent) from the “Manage Submittal Reporting Periods” area on the “Utilities” tab.</p>

Key	Control	Description
6	Process XML button	<p>Click the “Process XML” button” and the “Confirm Reporting Period – XML Input File Assignment” dialog appears.</p>  <p>If the file matches the Reporting Period, click “Yes”; otherwise click “No” and repeat the previous step.</p> <p>Follow the dialogs to confirm that the SBDD feature class is being properly associated with the Submittal Type.</p>  <p>Use the Submittal Type pull-down to select the matching Submittal Type for the current Feature Class Name. DIVA will attempt to find the best match for you by matching the feature class and Submittal Type names.</p> <p>Note: there will be one “Confirm Feature Class” dialog for each SBDD-defined feature class. The user will be asked to click “OK” twice, once to confirm the mapping as in the second illustration above and then again on a simple confirm dialog. It is faster to use the return key on the keyboard instead of clicking OK with the mouse.</p>

Key	Control	Description
7	View the XML Output button	<p>Click on the View the XML Output Button and the “XML Viewer” window opens:</p>  <p>Click to expand one node of the tree view; double click on a branch to expand all nodes nested below it. This is only used if you need to confirm something about the XML import.</p>
	“View Processing Log File” button	<p>Click on the “View Processing Log File” button and the “Log Viewer” window opens:</p>  <p>This is only used if you need to confirm something about the XML import process.</p>

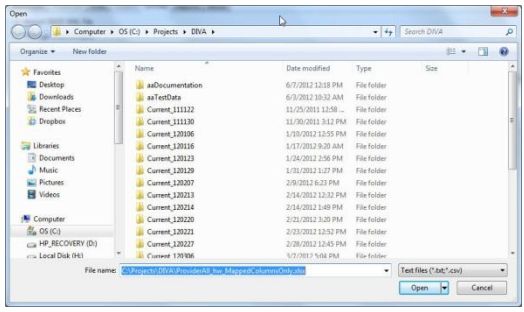
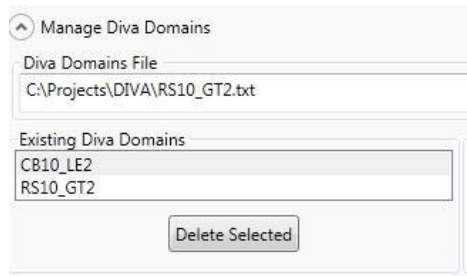
Manage DIVA Domains (Utilities tab, “Manage DIVA Domains” control)

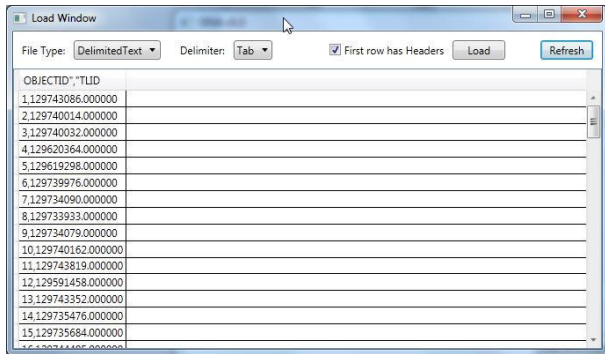
The “Manage DIVA Domains” control allows the user to import into a DIVA table a file (text, dbf or Excel) containing domain values. The imported domain table can then be referenced in a “Must Be in DIVA Domain” Rule from the “Rules” tab.

Click on “Manage DIVA Domains” section of Utilities Tab and the area expands

The screenshot shows the DIVA v5.2 application window. The 'Utilities' tab is selected in the top menu bar. The 'Manage Diva Domains' section is expanded, showing a 'Diva Domains File' text box with a 'Browse' button. Below this, there is a list of 'Existing Diva Domains' with 'RS2012_GT2' selected and a 'Delete Selected' button. To the right, the 'New Domain Info' section has a 'Domain Name' text box. Further right, the 'User Actions' section has a 'Load Data' button. At the bottom, there is a 'Status' section displaying: Provider Name: TW Telecom of Arizona LLC, Reporting Period: Spring 2012, Submittal Type: Service Address, and Input File Name: twtelecomofarizonallc_address_availability_AZ. A row of buttons at the bottom includes 'Help', 'Input Data Analysis', 'Output Data Review', 'Filter Records' (highlighted in red), and 'Edit Metadata'.

Key	Control	Description
1	“Diva Domain File” text box	Displays the path and file name to the SBDD XML file.

Key	Control	Description
2	Browse button	<p>Click on the “Browse” button and the browse dialog box opens:</p>  <p>Choose the file type filter at the bottom of the Browse dialog to match the file type (extension) of your domain file type.</p> <p>Select the source file and click “Open” on the file dialog box. The browse dialog goes away and you will be back to the Utilities tab with the pathname and file to your source file now showing in the “Diva Domain File” text box.</p>
3	Clear button	Clears the value in the “Diva Domain File” text box. Does not delete the source file or unload a previously loaded domain file. It just clears the text box.
4	Preview button	Opens the DIVA “Preview Window” which has been previously discussed.
5	Existing DIVA Domain area	<p>If the domain import was successful (Step 7), you will see the new domain name you provided in a previous step in the will appear under the “Existing DIVA Domains” area.</p>  <p>From here, you can select and delete and existing domain, as well</p>

Key	Control	Description
6	New Domain Info area	Assign the domain a name in the “Domain Name” text box and then load the data (see next step)
7	Load Data button	<p>The “Load Data” button becomes active after the user has selected a “Diva Domain File” in the previous steps and assigned a “Domain Name”.</p> <p>Click the “Load Data” button” and the “Load” window opens:</p>  <p>Like all DIVA Preview Windows, only the first 100 records of the source file will be shown.</p> <p>If the source is a text file, be sure to set the proper file delimiter in the “Delimiter” pull down menu at the top of the window.</p> <p>The “First row has headers” checkbox must be checked if the first row is column names.</p> <p>To populate the window with records from the source file or to se changes you made in this window, click the “Refresh” button in the upper right corner.</p> <p>To load the data, click the “Load” button at the top of the dialog.</p>

Tips for Managing DIVA Domains:

- Do not load large tables if avoidable. This will significantly slow performance when applying these domains through rules. So, for example, if you need to check for valid Census block ID values, you can

create and load a table based only on Census Blocks less-than-or-equal-to 2 square miles. There is no need to include all the larger Census Blocks in the domain table.

Update Provider Information (Utilities tab, “Update Provider” control)

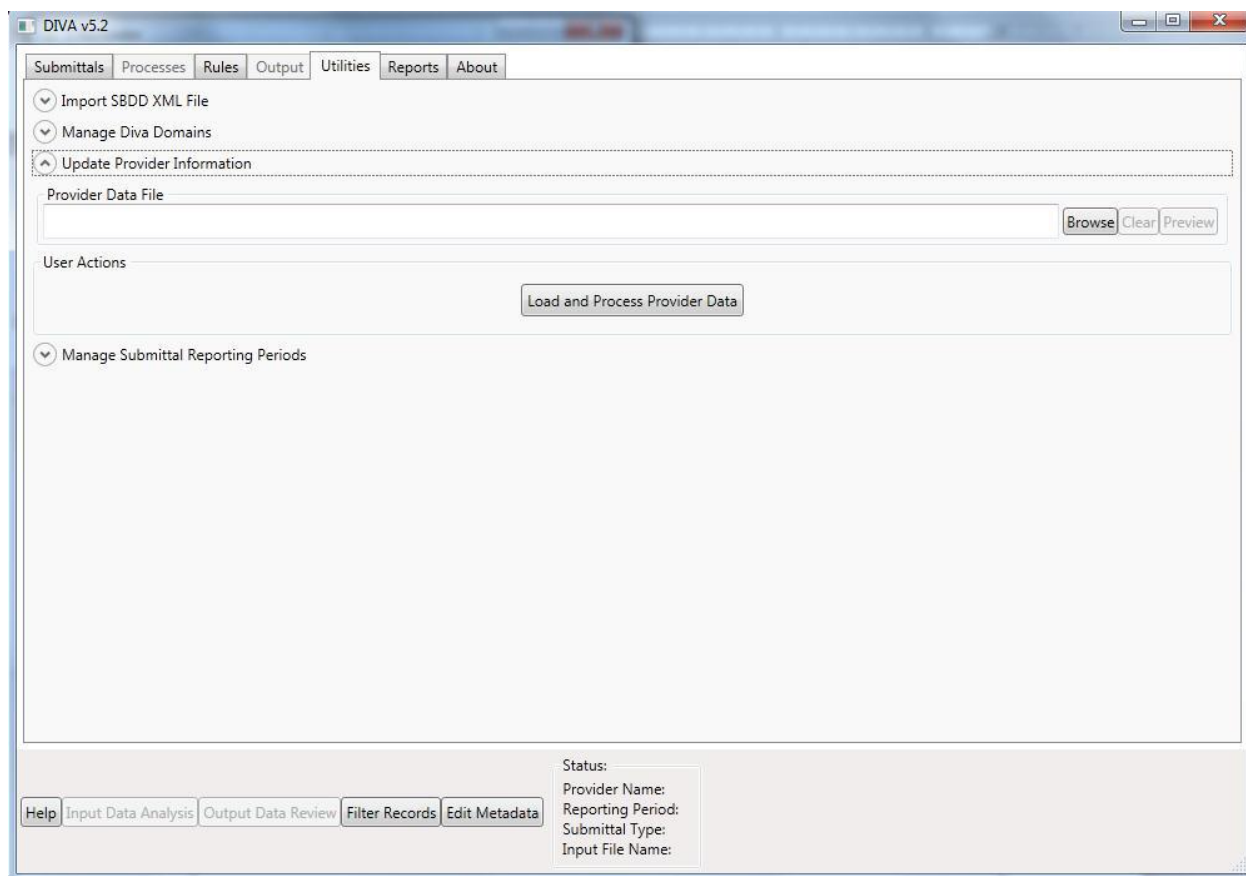
The “Update Provider Information” control allows the user to import a file (text, dbf or Excel) containing Provider identification and status information into a DIVA table. In the Arizona Broadband Project, this file is a spreadsheet maintained by the telecommunication consultants and uploaded to a wiki page. The information in the spreadsheet includes a two-character ProviderID value which is used internally to track Providers, as well as SBDD fields for Provider Name, DBA Name and FRN.

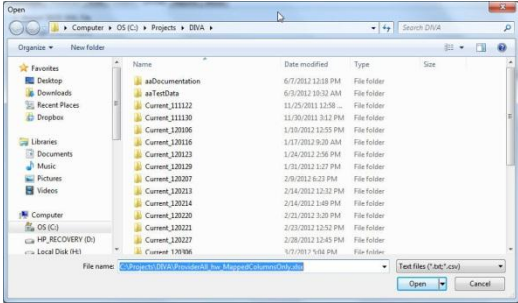
Users may need to pre-Process the file containing the new Provider Information. The source file must have a header row and the field names must match the following DIVA file names exactly (case and spaces, included).

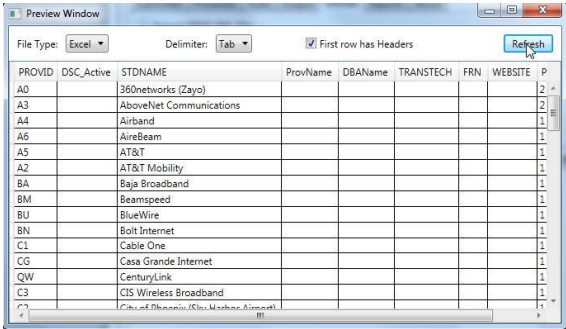
ProviderID	PROVNAME	DBAName
FRN	ProviderType	DSC_Active
Notes		

There can be additional fields or even missing fields (e.g. Notes), but DIVA must find an exact match on the header row for each column it imports.

Click on “Update Provider Information” section of Utilities Tab and the “Update Provider Information” area expands.

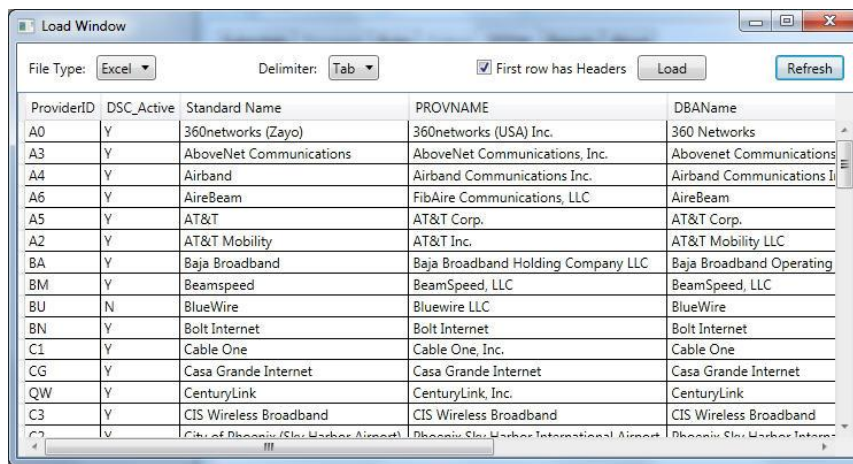


Key	Control	Description
1	"Provider Data File" text box	
2	Browse button	<p>Click on the "Browse" button and the browse dialog box opens:</p>  <p>Choose the file type filter at the bottom of the Browse dialog to match the file type (extension) of your Provider information file.</p> <p>Navigate to and select the source file and click "Open" on the file dialog box. The browse dialog goes away and you will be back to the Utilities tab with the pathname and file to your source file now showing in the "Provider Data File" text box.</p>
3	Clear button	Clears the value in the "Provider Data File" text box.. Does not delete the source file or unload a previously loaded Provider information file. It just clears the text box.

Key	Control	Description
4	Preview Button	<p>Click on “Preview” button and the “Preview Window” opens:</p>  <p>Like all DIVA Preview Windows, only the first 100 records of the source file will be shown.</p> <p>Be sure the “File Type” pull down reflects the extension of the source file.</p> <p>If the source is a text file, be sure to set the proper file delimiter in the “Delimiter” pull down menu at the top of the window.</p> <p>The “First row has headers” checkbox must be checked and the header row must have column names that exactly match what DIVA expects, or the import will not work. See the first step in this process for the list of valid header row names</p> <p>To populate the window with records from the source file, click the “Refresh” button in the upper right corner.</p> <p>Dismiss the “Preview” window by clicking the red “x” in the upper right corner of the dialog.</p>

Key	Control	Description
-----	---------	-------------

Click on the “Load and Process Provider Data” and the “Load Window” opens:



Like all DIVA Preview Windows, only the first 100 records of the source file will be shown.

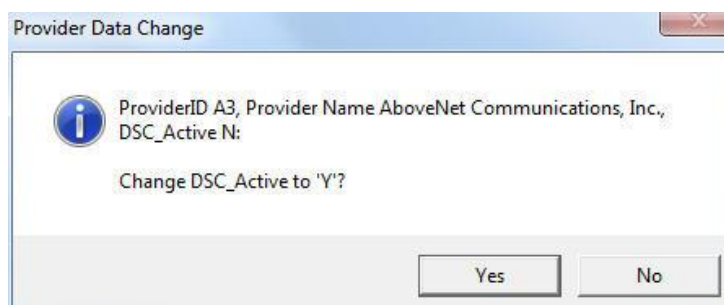
If the source is a text file, be sure to set the proper file delimiter in the “Delimiter” pull down menu at the top of the window.

The “First row has headers” checkbox must be checked and the header row must have column names that exactly match what DIVA expects, or the import will not work. See the first step in this process for the list of valid header row names

To populate the window with records from the source file, click the “Refresh” button in the upper right corner.

To load the data, click the “Load” button at the top of the dialog.

Assuming there are changes since the last version, a dialog will appear that asks you to confirm or reject each change since the last version.



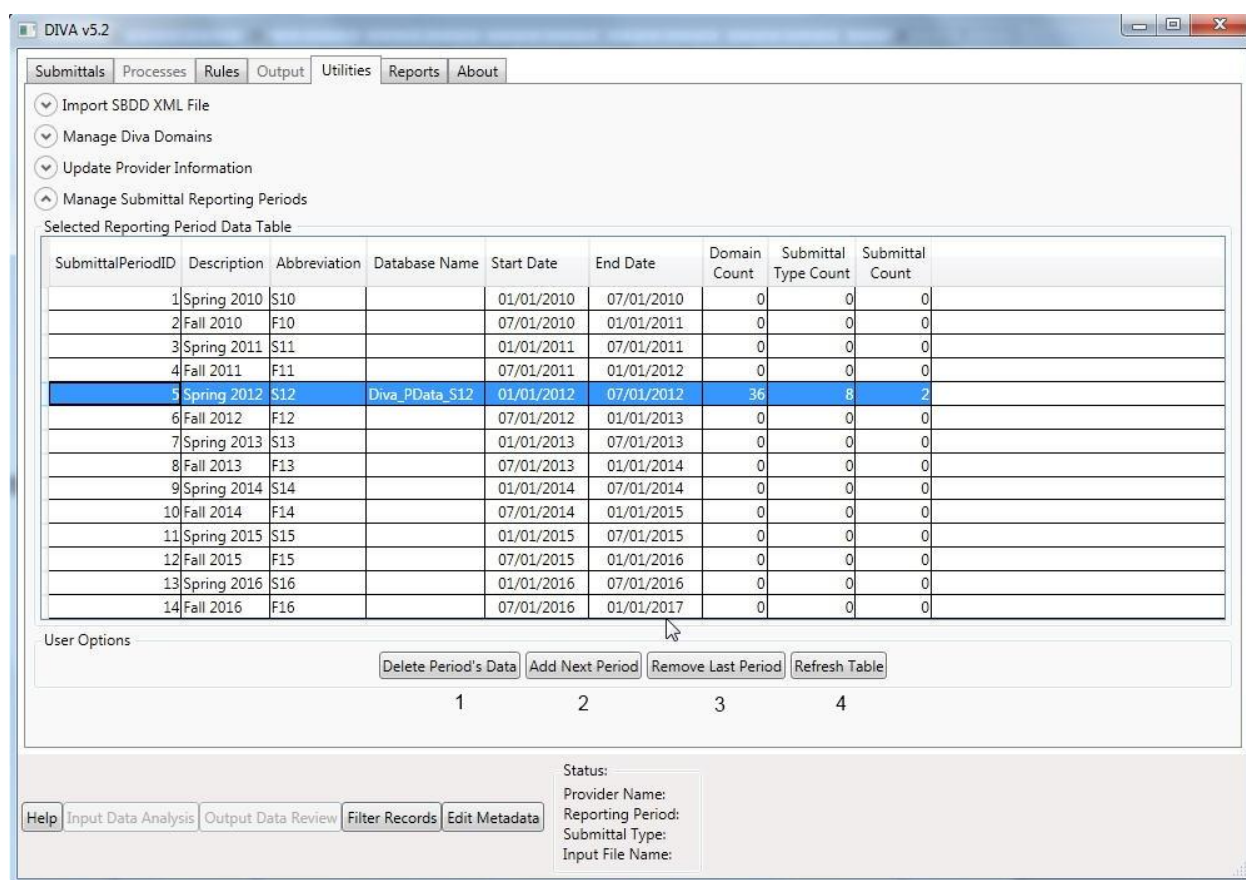
5

Load and
Process
Provider
Data button

Manage Submittal Reporting Periods (Utilities tab, “Manage Submittal Reporting Periods” control)

The “Manage Submittal Reporting Periods” control allows the user to add or delete Reporting Periods but not change any of the descriptive or date attributes.

Click on “Manage Submittal Reporting Periods” section of Utilities Tab and the “Manage Submittal Reporting Periods” area expands. Below are explanations of the “Manage Submittal Reporting Periods” control. Explanations of the buttons are keyed by number.



Key	Control	Description
1	Delete Period's Data button	Users can delete a Submittal Period's data which removes all database entries related to the Reporting Period.
2	Add Next Period button	Adds the next Submittal Reporting Period to the bottom of the table.

Key	Control	Description
3	Remove Last Period button	Removes the last Submittal Reporting Period from the bottom of the table.
4	Refresh Table button	Shows changes made to the table.

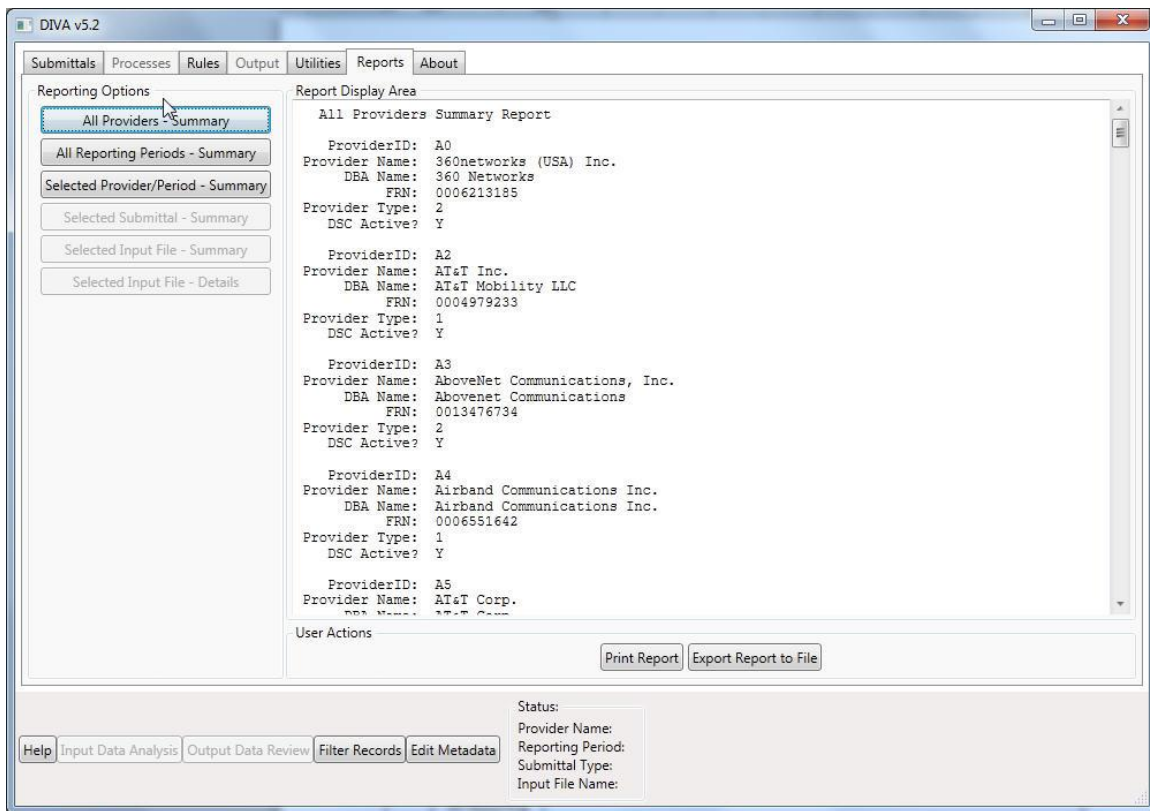
This control also shows some information about SBDD XML files that have been imported by the user. In the example above, only Spring 2012 SBDD information has been imported which entailed 36 domains covering 8 Submittal Types. In addition, the numbers of Submittals that have been created are displayed in the “Submittal Count” column.

Each Submittal Period has a 6-month window of time, though the Submittal Date is in the middle of that window. This allows changes to the Submittal Period data before the system clock closes access to a Submittal Period.

Reports Tab Functions

Reporting Option “All Providers - Summary” (Reports tab, button)

Creates a summary listing of current Provider information as imported into DIVA from the “Update Provider Information” utility.



Provider information consists of:

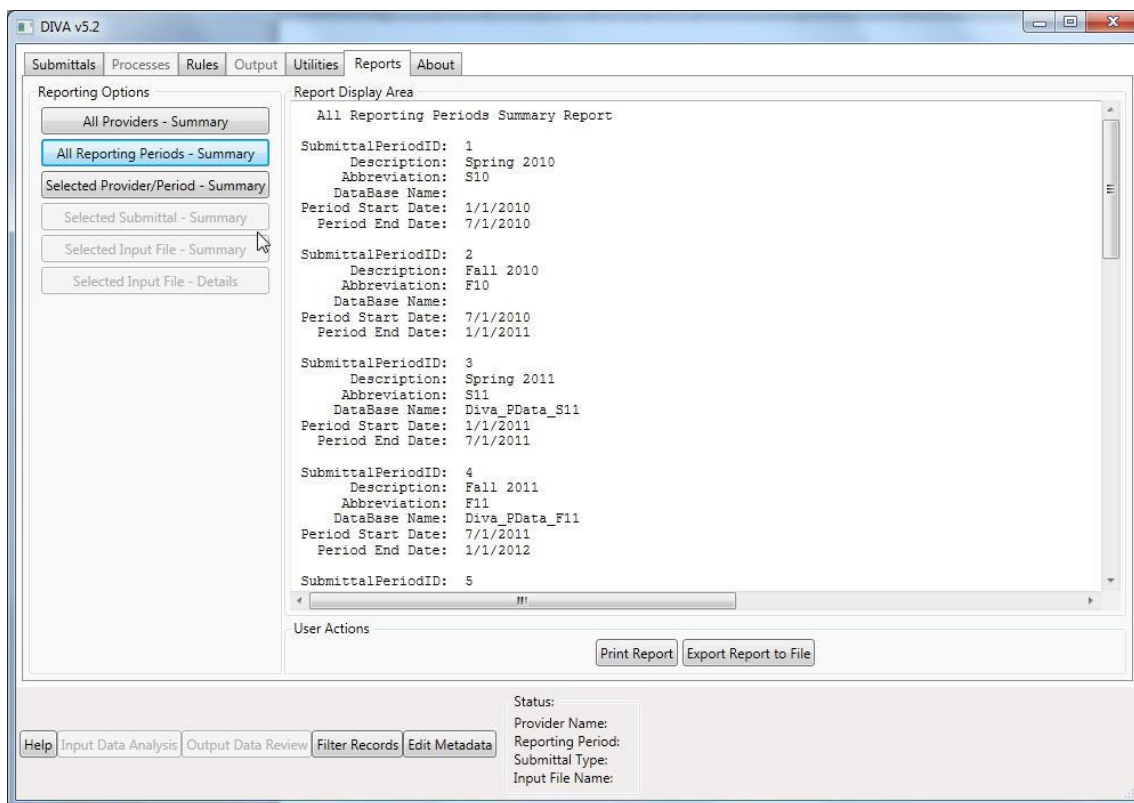
- ProviderID value – internal user-defined two-character identifier to uniquely identify Providers; it is not part of the SBDD database schema. This value, along with other Provider information, derives from imported Provider information. This value is used to create Provider-specific Input File data table names and Processing Set names.
- Provider Name value – the NTIA-required Provider Name attribute. This value is used in the first drop-down list in the Provider area on the Submittals tab.

- DBA Name value – the NTIA-required DBA Name attribute. This value is used in the second drop-down list in the Provider area on the Submittals tab.
- FRN – the NTIA-required FCC Registration Number attribute. This value is used in the third drop-down list in the Provider area on the Submittals tab.
- Provider Type value – the NTIA-required Provider Type attribute; this uses an SBDD-defined standard code.
- DSC_Active value – a Y / N entry that indicates whether this Provider is active in terms of providing data sets. This is a code developed by the AZBB team (not part of the SBDD database schema).
- Notes – user-added Provider-specific metadata (from Edit Metadata Window”); only appears if there is a note (not part of the SBDD schema)..
- Reporting Period Notes – user-added Reporting Period-specific metadata attached to this Provider (from Edit Metadata Window”); only appears if there is a note (not part of the SBDD schema)..

This set of data, except “Reporting Period Notes”, is repeated in other reports. When it is, rather than describing each reported field again, we will simply note “Provider information”.

Reporting Option “All Reporting Periods – Summary” (Reports tab, button)

Creates a summary listing of current Reporting Period information as currently stored in DIVA. The user may modify this using the “[Manage Submittal Reporting Periods](#)” utility. If a Reporting Period shows values for “Database Name”, it means that the SBDD geodatabase for that Period has been imported into DIVA using the “[Import SBDD XML File](#)” utility.



Reporting Period information consists of:

- Submittal Period ID – an internal identifier created by DIVA (not part of the SBDD schema).
- Description – a short description of the Submittal Period (not part of SBDD schema). This value is used in the Reporting Period drop-down list on the Submittals tab.
- Abbreviation – an abbreviated description of the Submittal Period (not part of the SBDD schema). This value is used to create Provider-specific Input File data table names and Processing Set names.
- Database Name – the DIVA name for the SQL Server database created for each Reporting Period using the “Importing SBDD XML File” button on the “Utilities” tab (not part of SBDD schema).
- Period Start Date – the starting date of the Reporting Period. These are defined based on the date of the data to be reported in the original NTIA reporting instructions (i.e. December 31 and June 30 of each year). The Reporting Period start date is the day after that date.

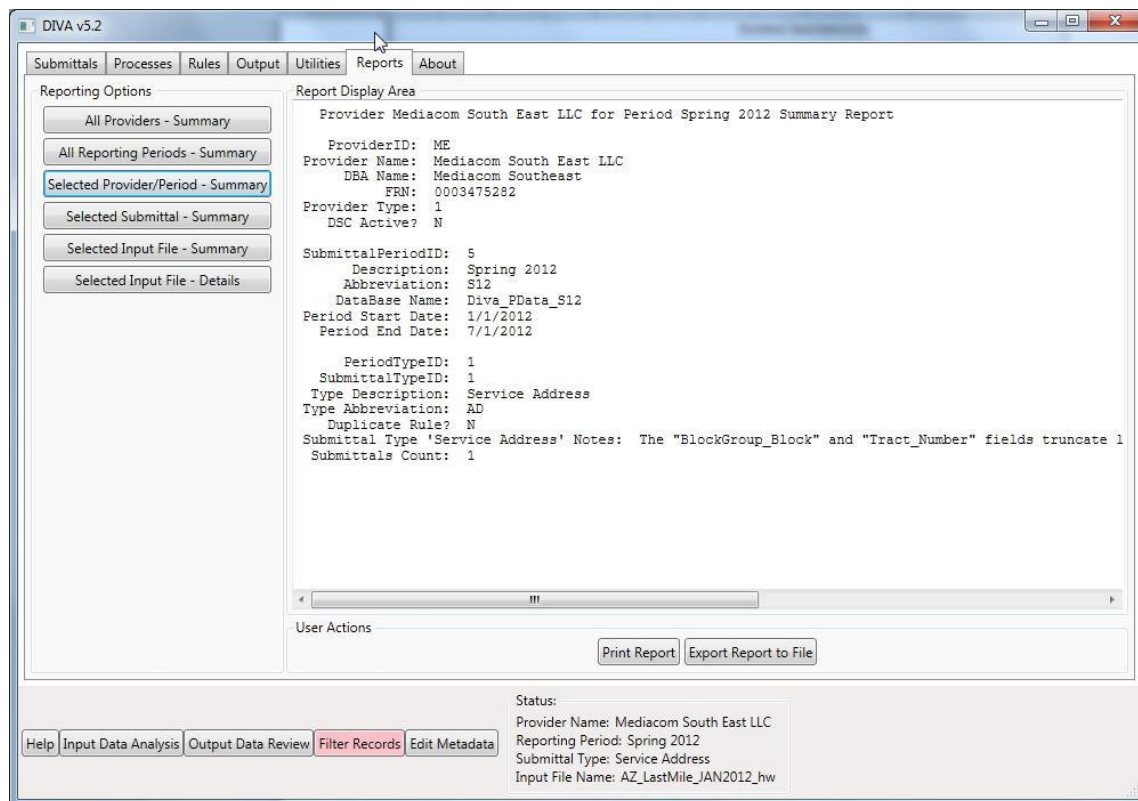
- Period End Date – the ending date of the Reporting Period. These are defined based on the date of the data to be reported in the original NTIA reporting instructions (i.e. December 31 and June 30 of each year). The Reporting Period end date is the day before the next Reporting Period’s Start Date.
- Notes – user-added Reporting Period-specific metadata (from Edit Metadata Window”); only appears if there is a note (not part of the SBDD schema).
- Provider Notes – user-added Provider-specific metadata attached to this Reporting Period (from Edit Metadata Window”); only appears if there is a note (not part of the SBDD schema).

Note that the Period End Date has nothing to do with the date that the NTIA requires information be submitted. The start and end dates are chosen so that there is no “gap” or “overlap” in the Reporting Periods.

This set of data is repeated in other reports. When it is, rather than describing each reported field again, we will simply note “Reporting Period information”. Note that the “Provider Notes” entry will be labeled differently so as to indicate both the Provider and Reporting Period.

Reporting Option “Selected Provider/Period – Summary” (Reports tab, button)

Creates a summary listing of information about the active Provider and the active Reporting Period.



Provider / Period information consists of:

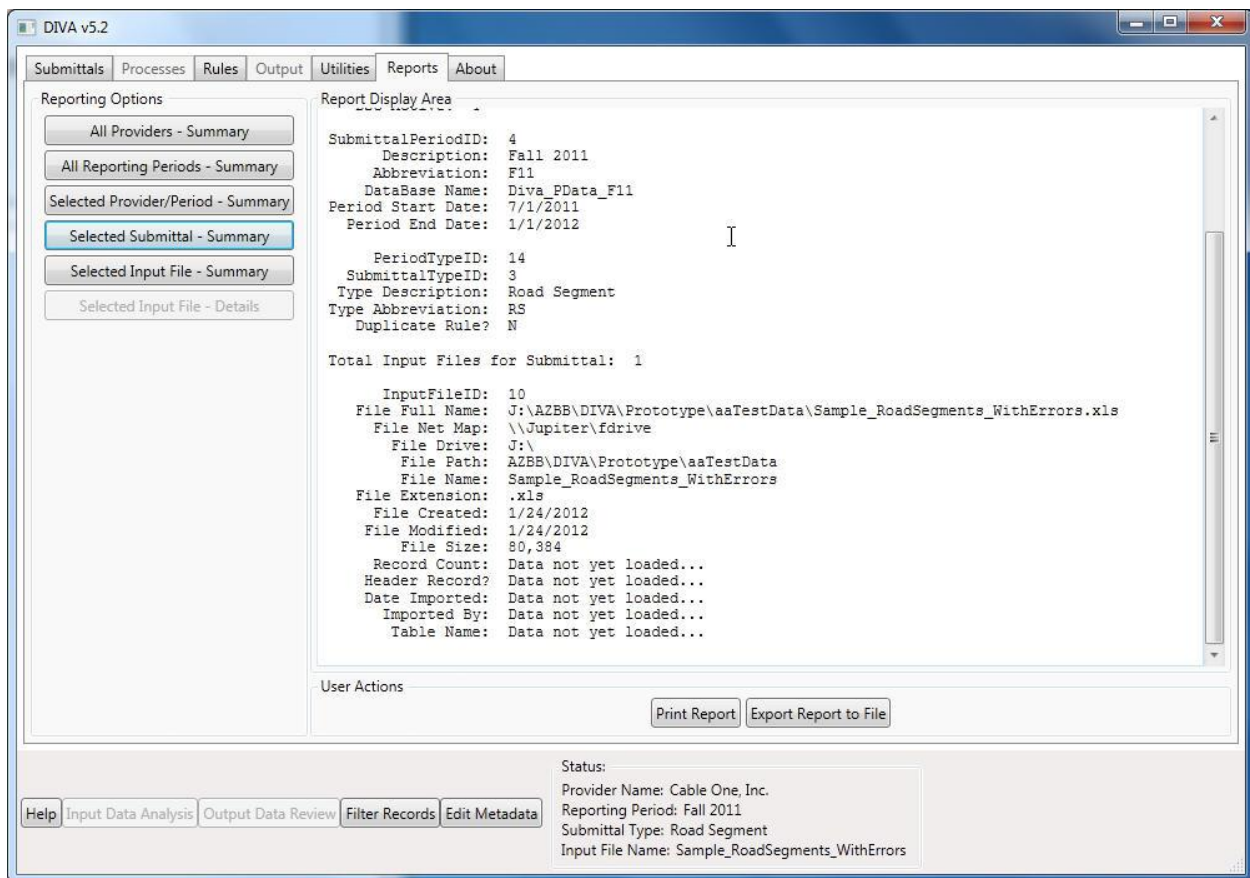
- Provider information.
- Reporting Period information.
- For each Period-Type in this Reporting Period (i.e. for each Submittal Type defined):
 - Period Type ID – an internal Period-Type identifier created by DIVA (not part of the SBDD schema).
 - Submittal Type ID – an internal Submittal Type identifier created by DIVA (not part of the SBDD schema).
 - Type Description – a short description of the Submittal Type (not part of SBDD schema). This value is used in the Submittal Type drop-down list on the Submittals tab.
 - Type Abbreviation – an abbreviated description of the Submittal Type (not part of the SBDD schema). This value is used to create Provider-specific Input File data table names and Processing Set names.
 - Duplicate Rule? – a Yes/No field indicating whether a “Must Not Have Duplicate” Rule has been assigned to any Template Output Field associated with this Period-Type. Used for DIVA internal processing (not part of the SBDD schema).

- Notes – user-added Period-Type-specific metadata (from Edit Metadata Window”); only appears if there is a note (not part of the SBDD schema).
- Submittals Count –the number of Submittals defined for the current Provider / Period-Type. Since there can only be one such Submittal (which may have several Input Files) for each Period-Type, this number will always be one. If there are none, then that Period-Type will not be reported (for brevity).

The set of data listed under Period-Type, except “Submittals Count”, is repeated in other reports. When it is, rather than describing each reported field again, we will simply note “Period-Type information”.

Reporting Option “Selected Submittal – Summary” (Reports tab, button)

Displays a summary of the active Submittal (button is inaccessible if no active Submittal). This report includes information about a single Submittal including a summary of the information about each of its Input Files, which includes record counts, the date imported, imported by etc. If the data for a Submittal has not yet been loaded, some of the Input File information will not be available (“Data not yet loaded...” is displayed as shown in the image below).



Submittal information consists of:

- Provider information.
- Reporting Period information.
- Period-Type information.
- Total Input Files for Submittal – the number of Input Files that are associated with this Submittal; the count must be at least one.
- For each Input File of this Submittal:
 - Input File ID – an internal Input File identifier created by DIVA (not part of the SBDD schema).

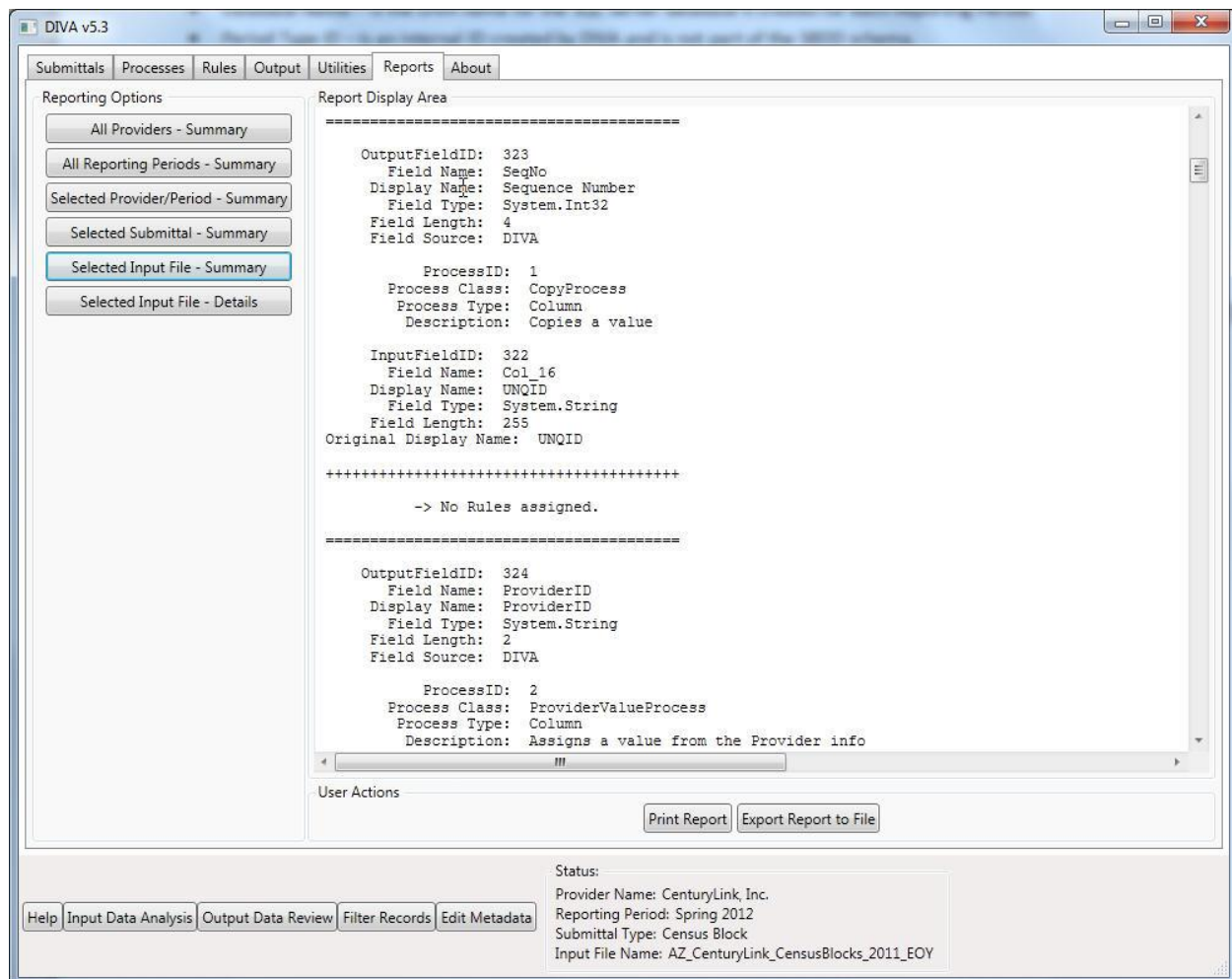
- File Full Name – the original complete file name, including path, of the Provider’s data set that was loaded into DIVA for processing (not part of the SBDD schema); this is the “source file”.
- File Net Map – if the drive letter portion of the File Full Name was a network mapping (share), then the value of that mapping is shown here (not part of the SBDD database schema).
- File Drive – the drive portion of the File Full Name (not part of the SBDD schema).
- File Path – the path portion of the File Full Name (not part of the SBDD schema).
- File Name – the file name portion of the File Full Name (not part of the SBDD schema).
- File Extension – the file name extension portion of the File Full Name (not part of the SBDD schema).
- File Name Comment – user-added File Name metadata (from Edit Metadata Window”); only appears if there is a comment (not part of the SBDD schema).
- File Created – the date that the source file was created (not part of the SBDD schema).
- File Modified – the date that the source file was last modified (not part of the SBDD schema).
- File Size – the size of the source file in bytes (not part of the SBDD schema).
- Record Count – the number of data records in the source file (not part of the SBDD schema). This is the number of records that are available for processing in DIVA; it excludes the count of the header record, if there was one.
- Header Record – a Y / N entry indicating if the data was loaded into DIVA with the “First row has Headers” checkbox checked. As explained in data loading, checking this box tells DIVA to use the values in the first record of the source file as the names of the Input Fields (data columns) and not to load it as data records for processing (not part of the SBDD schema).
- Header Record Comment – user-added Header Record metadata (from Edit Metadata Window”); only appears if there is a comment (not part of the SBDD schema).
- Date Imported – date that the source file’s data was loaded into DIVA (not part of the SBDD schema).
- Imported By – the Window’s user name of the person who loaded the data into DIVA (not part of the SBDD schema).
- Table Name - the name of the SQL data table and database that the data was loaded into (not part of the SBDD schema). This name is a combination of the Provider ID, Reporting Period’s Abbreviation, Submittal Type’s abbreviation and the Input File ID. See [Appendix D](#) for further details.
- Notes – user-added Input File-specific metadata (from Edit Metadata Window”); only appears if there is a note (not part of the SBDD schema).
- Projection Type – user-added Projection Type metadata (from Edit Metadata Window”); only appears if there is an entry (not part of the SBDD schema).
- Projection Note – user-added note regarding the projection type metadata (from Edit Metadata Window”); only appears if there is a note (not part of the SBDD schema).

The set of data listed under Input File is repeated in other reports. When it is, rather than describing each reported field again, we will simply note “Input File information”.

Reporting Option “Selected Input File – Summary” (Reports tab, button)

Creates a summary listing of the active Input File for the active Submittal (button is inaccessible if no active Input File). It adds Input File-specific information regarding the Output Fields of the Input File and the Process that is the source of data to each Output Field. It traces the “flow” of data “upstream”, through possible Process Fields and additional Processes until it reaches a data source, which is either an Input Field or a Process. It also lists all of the Rules that are attached to this Output Field along with a count of Rule violations. Accuracy in this report depends on the user having first visited the Processes tab and then the Output tab, clicking the “Refresh Violation Counts” button on that tab. Failure to do so will result in inaccurate reporting of error counts.

Since a Row Process (i.e., Resolve Duplicate Process) cannot be attached to any Output Field, the information regarding this Process is added at the end of the report.



Input File (summary) information consists of:

- Provider information.
- Reporting Period information.
- Period-Type information.
- Input File information.
- For each Output Field of the Input File (a double-dashed line separates each entry):
 - Output Field ID – an internal Output Field identifier created by DIVA (not part of the SBDD schema).
 - Field Name – the name of the field as specified in the SBDD schema. The Field Name is what is shown in the “tool tip” when you hover the mouse over the Output Field in the Processes tab.
 - Display Name – the alias of the field as specified in the SBDD schema.
 - Field Type– the field’s data type as specified in the SBDD schema but translated into a .NET data type.
 - Field Length– the field’s length as specified in the SBDD schema if the Field Type is a string; otherwise, the field’s length depends on the .NET data type.
 - Field Source – “SBDD” indicates that the field was specified in the SBDD schema; “DIVA” indicates that the field was added by the DIVA development team. Remember that all Output Fields derive from Template Output Fields and these are defined during the SBDD XML Import operation.
- For the Process connected to the Output Field (only one is possible):
 - Process ID – an internal Process identifier created by DIVA (not part of the SBDD schema).
 - Process Class – the type of the Process, e.g. “TranslationProcess” (not part of the SBDD schema).
 - Process Type – the type of the Process (not part of the SBDD schema). Only “Column” type Processes can be listed here.
 - Description– a short description of the Process (not part of the SBDD schema). The Description is what is shown in the “tool tip” when you hover the mouse over the Process in the Processes tab.
 - If the Process is a Constant Value Process:
 - Constant value set – the value that will be assigned to this Output Field for every record (not part of the SBDD schema).
 - Constant value type – the data type assigned to the value during user-configuration of the Process (not part of the SBDD schema).
 - Because this Process serves as a data source, no further “tracing” to find additional Processes is done.
 - If the Process is a Provider Value Process:
 - Provider attribute – the property (attribute) of the Provider (e.g., “ProviderType”) whose value will be assigned to this Output Field (not part of the SBDD schema). This is set during user-configuration of the Process.
 - Provider value set – the value of the Provider’s attribute that will be assigned to this Output Field for every record (not part of the SBDD schema).
 - Provider value type – the data type of the specified Provider attribute (not part of the SBDD schema).
 - Because this Process serves as a data source, no further “tracing” to find additional Processes is done.

- The data source(s) for each Process may be either a Process Field or an Input Field. If there are multiple sources, they will be listed consecutively (although they may be interspersed with Processes, etc., as the data flow is “traced” from Output Field back to Input Field)
- For each connected Process Field:
 - Process Field ID – an internal Process Field identifier created by DIVA (not part of the SBDD schema).
 - Field Name – the name of the field as assigned by DIVA (not part of the SBDD schema).
 - Display Name – the display name of the field as assigned by DIVA (not part of the SBDD schema). The Display Name is what is shown in the “tool tip” when you hover the mouse over the Process Field in the Processes tab.
 - Field Type– the field’s data type (not part of the SBDD schema). Typically this is left blank.
 - Field Length– the field’s length (not part of the SBDD schema). Typically this is left blank.
 - The Process connected to the Process Field (only one is possible) and any further elements as data flow are “traced” back to source(s).
- For each connected Input Field:
 - Input Field ID – an internal Input Field identifier created by DIVA (not part of the SBDD schema).
 - Field Name – the name of the field as assigned by DIVA (not part of the SBDD schema). This will be either “InpSeqNo” for the field that holds the sequential record number or a name of the form “Col_##” where “##” is a two digit integer starting at 0 (e.g., “00”, “01”, etc.). Field names are automatically assigned by DIVA during data loading depending on the position of columns in the user’s data source file.
 - Display Name – the display name of the field as assigned by DIVA (not part of the SBDD schema). The Display Name is what is shown in all of the displays of Input Fields. It will be the same as the Field Name if no header record was specified during data loading or the value found in the header record for that field or a user-specified value (set in Input Data Analysis window or in the Input Fields list on the Processes tab).
 - Field Type– the field’s data type (not part of the SBDD schema). This will always be string except for the “InpSeqNo” field, in which case it is Int32.
 - Field Length– the field’s length (not part of the SBDD schema). This will always be 255 except for the “InpSeqNo” field, in which case it is 4.
 - Original Display Name – the display name of the field as originally assigned by DIVA (not part of the SBDD schema). It will be the same as the Field Name if no header record was specified during data loading or the value found in the header record for that field in the header record.
 - An Input Field is a data source so tracing stops at this point (although there may be additional Input Fields to list, or previous “tracings” to complete).
- For each Rule attached to an Output Field (Rule listings are separated from the first part of the Output Field listing by a row of plus signs):
 - Rule ID – an internal Rule identifier created by DIVA (not part of the SBDD schema).
 - Rule Type – the type of Rule (not part of the SBDD schema). This will always be “Column” for all rules listed with an Output Field.

- Rule Description – the description shown in the Rules tab along with the Rule name (not part of the SBDD schema). The Description is what is shown in the “tool tip” when you hover the mouse over a Rule violation count in the Output tab.
- Rule Source– the source for the Rule. It will be “SBDD” if it was created as part of the SBDD XML Importoperation or “User” if it was assigned by the user on the Rules tab.
- Violation Count – the number of errors for (violations of) this Rule since the last “Refresh Violations Count” button click. A value of “-2” indicates that the Rule has not been configured (for user-assigned Rules); a value of “-1” indicates that there is no data to check (no data source attached to the Output Field).

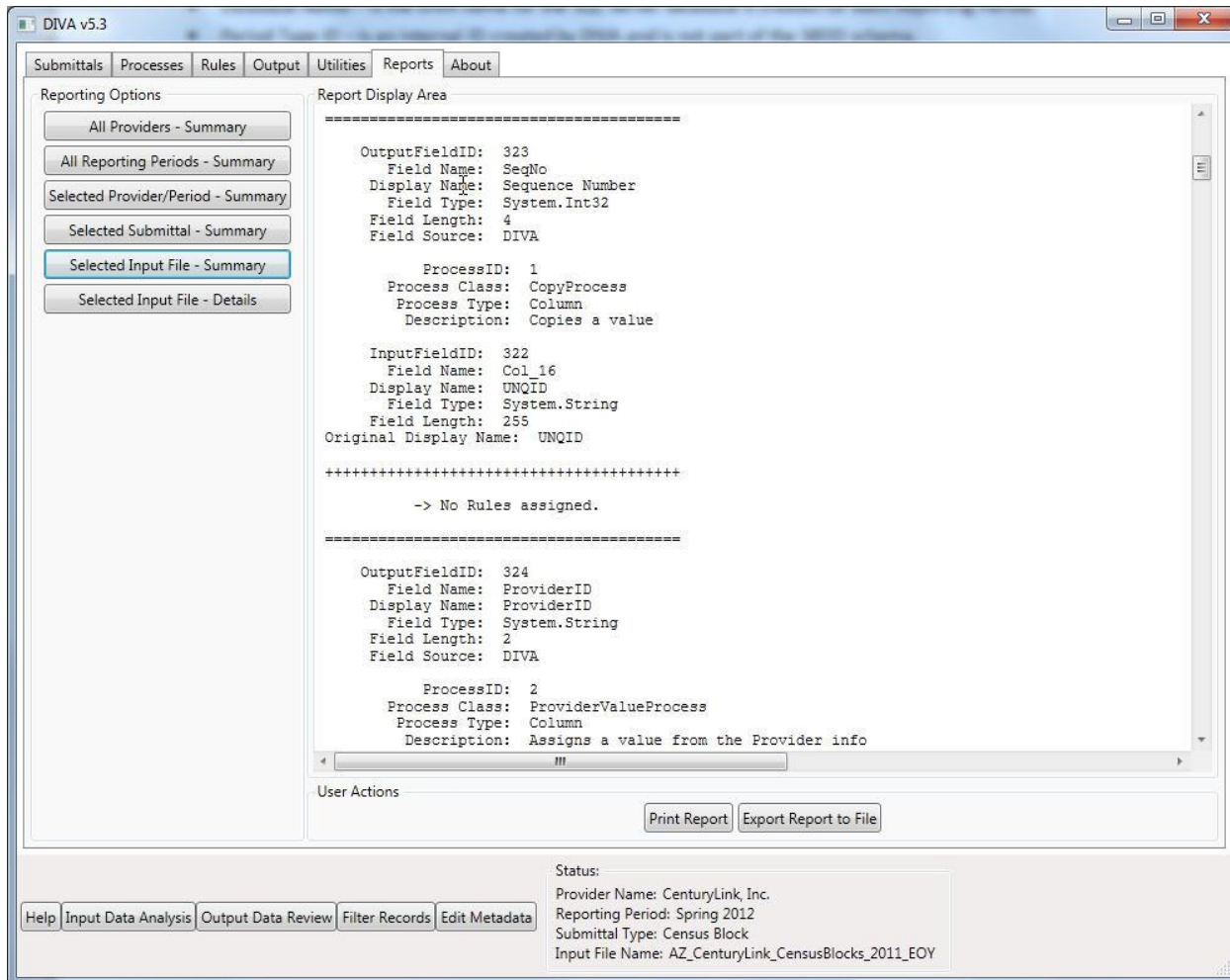
- The set of data listed after Output Field is repeated in the “details” version of this report. It will be referenced there as “Output Field information”.

- The final entries after all Output Fields have been listed are the Row Processes associated with Row Rules. There is currently only one Row Rule (“MustNotHaveDuplicate”), so there is only one Row Process (“ResolveDuplicate”) to list. This will only be present if this Rule has been assigned on the Rules tab to this Period-Type.
 - Process ID – an internal Process identifier created by DIVA (not part of the SBDD schema).
 - Process Class – the type of the Process, e.g. “ResolveDuplicateProcess” (not part of the SBDD schema).
 - Process Type – the type of the Process (not part of the SBDD schema). Only “Row” type Processes can be listed here.
 - Description– a short description of the Process (not part of the SBDD schema).
 - Decider Field #1 – the first Output Field used to determine if a duplicate exists.
 - Decider Field #2 – the second Output Field used to determine if a duplicate exists.
 - Repeated for each Output Field that is having its value resolved in the case of duplicates existing is a statement giving the name of the Output Field and how the resolution is to be made (currently either “Max value” or “Min value”).

The set of data listed above is repeated in the “details” version of this report. It will be referenced there as “Resolve Duplicate information”.

Reporting Option “Selected Input File – Details” (Reports tab, button)

Creates a summary listing of the active Input File for the active Submittal (button is inaccessible if no active Input File). It adds some additional detailed information to the “Input File – Summary” report discussed previously.



Input File (details) information consists of:

- Provider information.
- Reporting Period information.
- Period-Type information.
- Input File information.
- Output Field information augmented with (additional information is bracketed with single-dashed lines):
 - Translation values – under the Translation Process, if present, the detailed “From” and “To” values (user-assigned in the configuration dialog box) are listed.
 - Rule violations – under each Rule that has a violation count greater than 0 will be listed each violating record. This essentially is a listing of what appears in the “Error Record View” panel on the Output tab.
- Resolve Duplicate information.

Sample Workflow

DIVA is an event driven application meaning that the user can control in large part what action is performed and when; therefore, the workflow below is only one of many possible paths for processing a file. This workflow is grouped into actions that are non-Reporting-Period specific, and Period specific. Within the Period specific workflow are General Operations (generally performed once) and Provider Operations (performed one-to-many times for each Provider). This workflow assumes you have opened DIVA and are ready to begin processing for a new Reporting Period.

Non-Reporting-Period Specific

1. [Import Provider Information](#) – only if there have been changes since the last Reporting Period. It is important that the ProviderID never change from version to version of the Provider Information table.
2. [Import DIVA Domains](#) – only if you have a need to validate values in a field that is not already covered in one of the imported SBDD or DIVA domains. For example, you may want to check each Provider's County FIPS codes against a table of valid codes in your State.

Reporting Period Specific – General Operations

3. [Maintain Reporting Periods](#) – the only time this will be necessary is when the standard list of Reporting Periods is exhausted and you need to add another one to the end of the list.
4. [Import the SBDD XML Files](#) - this is how DIVA enforces the SBDD rules built into the domains in the SBDD database for each Reporting Period. THIS STEP MUST BE PERFORMED FOR EACH REPORTING PERIOD FOR WHICH DATA WILL BE PROCESSED. If you are not sure which Reporting Periods have been imported, you can use the “[Manage Submittal Reporting Periods](#)” function on the Utilities tab to see which ones have a value in the “Database” column. Note that DIVA will not let you import a given Reporting Period twice.
5. [Assign/Modify/Delete Rules](#) – the SBDD import function will automatically assign Rules to the Template Output Fields for a Period-Type which means there will already be SBDD-defined rules that will attach to every Output Field for every Submittal type you will process. You can assign your own Rules or even delete the imported SBDD rules. Be careful though. Once you delete a rule for a Period-Type it is deleted for every subsequent Provider for that the Period-Type. It also means any Provider you have processed prior to the delete was checked against that rule but any going forward will not be. This goes for adding new rules after one or more Providers have already been processed. You may want to avoid this inconsistency.

Reporting Period Specific – Provider Operations

6. [Pre-process Provider Input Files](#) – DIVA is designed to handle Input Files in a row –column structure and the user must verify and modify them to meet this criteria.
7. Import a Provider Input File –
 - a. [Select an Input File](#) – browse to and select an Input File for loading.
 - b. [Preview the selected Input File](#). – you do this as the very first step to see what you have.
 - c. [Filter Records](#) - as needed based on evaluation from the Input File preview.
 - d. [Load the selected Input File](#) – loads data from Input File into an internal DIVA table.
8. [Analyze the Loaded Data](#) – use DIVA tools to better understand the data in the Input File and help you make choices for subsequent processing.
9. [Define and Save Processes](#) – apply transformation Processes to values from the Input Fields to values in the Output Fields. On the “Processes” tab
 - a. Be sure to use a [“Copy” Process](#) to save the value of DIVA input field, “InpSeqNo” to DIVA Output field, “SeqNo”. This will give you a way to tie Input and Output records together. In the case of shapefile processing, use “SeqNo” to capture a unique identifier from the Input file for later use in joining the clean DIVA text file back to the shape.
 - b. Use [“Provider Value”](#) processes to populate the provider identification Output Fields. This will help ensure that the final NTIA deliverable is consistent with the official project view of provider identification. Do note in the metadata (later to be copied to the QA spreadsheet) if there is a discrepancy between the Provider-submitted values and the ones that result from the “Provider Value” process. You can quickly spot errors by comparing the values in the Input Field list with those that show up in the Output Field list after Processes are connected and applied.
 - c. Be sure to “Apply” and “Save” your Processes regularly so you do not lose your work.
 - d. Do not hesitate to use the [“Input Data Analysis”](#) button from the Status Bar any time you need to better understand values in Input Fields and therefore the correct process to apply to transform these values to Output Fields. For example:
 - i. The Input Field “TransTech” may be displaying a valid NTIA code, “10” but you are not sure if every record in the Input File is a “10.” Open the “Input Data Analysis” window and run a “Do Frequency” to see every unique value for “TransTech” in the Input File.
 - ii. You are not sure if the Census Block identifier is unique or whether there are multiple records for some Census blocks. A “Do Frequency” will give you a “Count” field on which you can sort to see if there are any values greater than 1.
 - e. If this set of processes seems generic enough to be applied to the same Submittal Type from subsequent Providers in this Reporting Period, then click the “Save as a Template” button.

10. [Load, Evaluate and Modify the Rules](#) – use the “Rules” tab to “Load”, inspect and edit as needed the rules that will be applied to the current Submittal. In this sample workflow, we suggest you do this step early on; however, after working with the data through the establishment of Processes, you may gain insight into new Rules that should be applied.
 - a. On the Rules tab, choose the “Reporting Period” and “Submittal Type” from the pull down menus that correspond to the active Submittal and click the “Load Rules” button.
 - b. [Evaluate the Currently Assigned Rules](#) - double click on the topmost Output Field to expand each Output Field to reveal the Rules nested under each one. Double click on any Rule you wish to evaluate and/or edit; right-click and select “Delete” from the context menu to delete a Rule.
 - c. [Add a new Rule](#) by dragging and dropping a Rule Type onto a field in the “Currently Assigned Rules” area and then configuring it appropriately.
11. [Apply the Rules and Look for Errors](#) – use the “Output” tab to see where Output Field values are not congruent with the Rules.
 - a. Click the “Refresh Violation Counts” button – depending on the number of and type of Rules, and the number of records in the Input File, this may take 20 to 30 seconds or more to complete.
 - b. [Use the “Rules Violations and Counts” area](#) to evaluate the result of applying Processes on the Output Fields - double-click the topmost Output Field in the “Rules Violations and Counts” area to reveal each Rule and the error count immediately to the right of each Rule.
 - c. [Evaluate the errors](#) - click on each Rule that has an error count greater than zero. Details on the error will appear in the “Failed Record View” area.
 - d. Based on your evaluation of errors, adjust the Filters, Export Options, or Processes and add, delete or modify Rules to eliminate or reduce the number of errors.
12. [Use the Output Data Review Window](#) – use the controls on this window to sort and filter errors and gain insight into why errors are occurring.
 - a. Adjust the Filters, Export Options, or Processes and add, delete or modify Rules to eliminate or reduce the number of errors.
 - b. Export the Data – to a delimited text file format. From here you can process the data onto the GIS layers for creation of the NITA submittals.
13. [Create/Edit Metadata](#) – click on the “Edit Metadata” button in the “Status Bar” area at any time to make notes about Reporting Periods, Providers, Provider Types, and more. This data can be output to reports and used to help fill in the QA Sheets used for each Provider
14. [Generate Reports](#) – click on a report type button to display various levels of detail for the current Reporting Period, Provider, Submittal Type, etc. Information from these reports, including metadata the user has captured using the “Edit Metadata” button, can be cut and pasted into the Provider QA worksheet.

Notes on Desired Functionality and Bugs

As with any software project, not all design targets were fully achieved or met in the way we anticipated. In addition, the process of developing DIVA gave us some good ideas for functionality that was not in the original project scope or budget. Fortunately, DIVA provides a framework on which additional functionality can be added. The source code is available to the BB community and is well-documented.

Things that don't and won't work:

- Compatibility with 32-bit OS: The specification in the original User Needs was that DIVA run in a 32-bit Windows operating system (Win 7 or XP). However, very large source files (1 million records or so) cannot be processed as DIVA does all data processing in memory (for speed) and 32-bit versions of the OS run out of addressable memory space.

Some ideas for additional functionality:

- Add a “drill-down” capability to the Output Data Review Window so that a user can see all of the errors affecting a record. The Output tab shows the errors one Rule at a time, making it difficult to see that one record may be violating multiple rules.
- Add some “intelligence” to the Process Mapping canvas so that graphics gracefully respond to re-sizing the DIVA GUI, adjusting their layout in a “proportional” manner.
- Right-click, copy and paste a process in the Process Mapping Canvas, rather than having to drag a new process from the Process Palette each time.
- More fully develop the “Full address decomposition” Process. This was not included in the current version due to budget priorities. And Arizona does not submit address data to NTIA per the original NDA's signed with each Provider.
- Modify the Object Data Model (ODM) so that Providers are treated like Submittal Types. By that we mean create a Period-Provider entity so that the “state” of each Provider for each Reporting Period could be retained. This would prevent some problems with users changing Provider ID values between Reporting Periods and allow us to track the history of Provider information changes.
- Sort columns as numbers, rather than text. You will notice that when you sort columns in any of the Preview or Analysis windows, the values sort as text. This is because DIVA treats all data as strings. Rules that check if data is an integer type, for example, do this without changing the data from a string.

No non-trivial programs are ever error free. We have done extensive testing of DIVA, but no doubt bugs remain. Please inform us of any that you might encounter.

Please share code enhancements with us and the DIVA community.

Glossary

active – the currently selected attributes: Provider, Reporting Period, Submittal Type, Submittal and / or Input File. An existing Submittal is made active by clicking on it in the “Existing Submittals” list in the Submittals tab (this list is the table on the bottom panel of the Submittals tab); this Submittal will then be highlighted in that list. Activating a Submittal will synchronize the drop-down lists in the upper panel of the Submittals tab with this selection; thus the active Provider, Reporting Period and Submittal Type will be shown. The first Input File of the Submittal will become active; the Input File drop-down list will be synchronized to show that Input File. Making any choice in the upper panel active (selecting a Provider, Reporting Period and / or Submittal Type) will activate the matching Submittal if it exists. If it does not exist, then a new Submittal may be created for it once an Input File is selected. Creating the Submittal (with the “New” button), will create it and make that Submittal active. If a Submittal has more than one Input File associated with it, a different Input File may be made active (using the Input File drop-down list) without changing the active status of any of the other attributes. Thus “active” is synonymous with “selected”.

Column (or Data Column) – a particular reported, or added, attribute of a data set when viewed as a data Table. The definition of the data values in this column remains constant throughout the Table, although the actual values vary by the Record or Row.

Column Process – a Process that is applied to all values in a particular data column (data source Field). All Processes that are manipulated on the Process Mapping Canvas are Column Processes.

Column Rule – a Rule that applies to all values in a particular data table column, that is, to an Output Field in all rows of the data table. Column Rule violations are corrected by adjusting Processes (creating, linking, configuration, removing) on the Process Tab.

Data Set – a file containing Provider-specific information, submitted to the State during a Reporting Period. Typically, this data set will be a text file, a spreadsheet file, or an ESRI Shapefile. DIVA requires that data sets be in a table-oriented format and handles only specific types of delimited text formats, Excel spreadsheets (in either .xls or .xlsx formats), or DBF formatted data tables. Provider submittals that do not meet formatting requirements require some pre-processing; see [Appendix B](#) for details.

Data Sink – a Field that serves as a target for the output of a Process. All Output Fields serve as data sinks. Process Fields that connect Processes may not serve as a data sink since, by definition, they are allowed only

one connecting line (one connection leading in) and that line is already assigned by the creation of a Process Field from an Input Process Field and an Output Process Field.

Data Source – a Field that serves as the source of data (input) for a Process. All Input Fields serve as data sources. Process Fields that connect Processes also serve as data sources for the Processes they feed into.

Data Validation – the process of checking a data value against another data value. Validation implies that knowledge outside of that data value is needed, that is, that another data source is involved. For example, checking a data value against a Domain of allowed values is a form of data validation.

Data Verification – the process of checking a data value without using another data value. Verification implies that knowledge outside of the checked data value is not needed, that is, that what needs checking is intrinsic to the data value itself. For example, checking that a data value is of a particular type (e.g., an integer value) does not require checking another data value and is thus a form of data verification.

Default Value – a value for an Output Field that will be entered if a Provider fails to report (specify) a value. Default values are specified in the SBDD Geodatabase.

DIVA Domain – a domain imported by the DIVA user which is not part of the SBDD database domains. An example of a DIVA Domain might be a list of valid Census Block ID's. This is not specified in the SBDD data model, but is useful for validating a Provider's list of Census Blocks. Users import DIVA Domains from the "Utilities" tab.

Keyed Domain – a domain whose scope is determined by the value in a key field. In the latest SBDD model, the speed fields are in a domain keyed off of the value contained in the Technology of Transmission field. This means that each Technology Type will have its own range of permissible speed values.

Domain – a set of values that are used to validate data.

Field – an entity that can contain provider reported data. A Field is always of some particular type: Input Field, Template Output Field, Output Field, Process Field, Input Process Field or Output Process Field.

Input Field – a data source for DIVA processing. The Input Fields for a particular Input File are created, and their values populated, based on the data fields (or “columns”) in a Provider’s submitted data set (the Input File).

Input File – the internal DIVA equivalent of a Provider-submitted [Data Set](#). An Input File, once associated with a Submittal, still needs to be “loaded”, that is, have the data set’s data values transferred into the internal storage database table that will be used for all processing. The actual Provider’s data set is never altered by DIVA.

Input Process Field – an “unconnected” [Data Source](#) Field for a Process. These will show up as one or more orange half circles with the arc facing left on the left side of a Process when it is dragged onto the Process Canvas. Not all Processes will have an Input Process Field(s).

Output Field – a target field ([Data Sink](#)) for DIVA processing. The Output Fields are either specified by the SBDD geodatabase for a given Reporting Period or added by DIVA designers to hold information that may not be part of a final SBDD feature class. This additional, non-SBDD information was deemed critical by the DIVA designers to post-DIVA GIS processing. Output Fields have their data values populated by transforming Input Field values using the Processes that a DIVA users specifies.

Output Process Field – an “unconnected” data sink Field for a Process. These will show up as one or more orange half circles with the arc facing right on the right side of a Process when it is dragged onto the Process Canvas. All Processes will have an Output Process Field(s).

Period-Type – a composite of the Reporting Period and Submittal Type. Because the reporting requirements for a particular Submittal Type may change each Reporting Period, this composite allows us to attach those specific requirements (attributes) to this particular Reporting Period / Submittal Type combination.

Process – an operation in DIVA that transforms data from an Input Field to an Output Field. DIVA has eight Processes that a user can apply to data. These are “Constant Value”, “Provider Value”, “Copy”, “Copy with Default”, “Make Negative”, “Translation”, “Composition”, and “Decomposition”. Each of these Processes is explained in detail in the “Functional Detail” section of this documentation.

Process Field – an intermediate data “storage” that DIVA uses to link (couple) Processes together when the connection is not directly to an Input or Output Field.

Processing Set – a collection of Processes defined for a particular Submittal. Processes are defined on the Processes tab of the DIVA GUI. A Processing Set is specific to a Provider, Reporting Period, Submittal Type and an Input File

Provider – a Broadband services Provider as defined by the NTIA. These may be primary providers or resellers.

Record (or Data Record) – see [Row](#).

Reporting Period – an NTIA established reporting period for submittal of State broadband data. There are two reporting periods per year, Fall (October 1st) and Spring (April 1st). In DIVA, there is a 6-month window assigned to each of these periods. The Fall reporting period spans July 1 to December 31 while the Spring period runs from January 1 to June 31. Only the Reporting Period that has a date range matching the current system date will be considered. The "Reporting Period" pull down list can be modified using the "[Manage Submittal Reporting Periods](#)" function found on the "Utilities" tab.

Row (or Data Row) – a particular instance of all reported, or added, attribute values of a data set when viewed as a data Table. The definition of the data values in this Row is set by the definition of the Columns of the Table. Each Row corresponds with a single observation of the values for the defined Columns and the order that the values have in the Row corresponds with the order of the defined Columns. A (Data) Record is synonymous with a (Data) Row.

Row Process – a Process that is applied to certain user-assigned Output Field values on particular data rows. The nature of the Row Process defines how particular data row groups are obtained, and the user-assigned Output Fields have their values determined from that particular data row group. There is, at present, only one Row Rule in DIVA ("Must Not Have Duplicate").

Row Rule – a Rule that applies to values in more than one column (Output Fields) across two or more rows of the table. Row Rule violations are corrected by configuring a specific Process (which is not user-creatable; it is created automatically by the existence of the Row Rule) on the Process Tab.

Rule – a Rule expresses a criterion that an Output Field's values must meet in order not to be flagged as an error (a Rule violation). DIVA has 10 pre-defined Rules that are classified according to their data Verification or Validation type, and their Column or Row type. These rules, in alphabetical order, are:

- MustBeInDivaDomain: Validation, Column.
- MustBeInDomain: Validation, Column.
- MustBeInKeyedDomain: Validation, Column.
- MustBeOfType: Verification, Column.
- MustHaveData: Verification, Column.
- MustNotHaveDuplicate: Validation, Row.
- NotGreaterThanField: Validation, Column.
- NotGreaterThanValue: Validation, Column.
- NotLessThanValue: Validation, Column.
- NotLongerThan: Verification, Column.

Rules are automatically created during the SBDD XML Import operation for a given Reporting Period (for all Submittal Types of that Period). These Rules can be reviewed and modified (or removed), and additional Rules assigned by users (from the “Rules Palette”) on the “Rules Tab”.

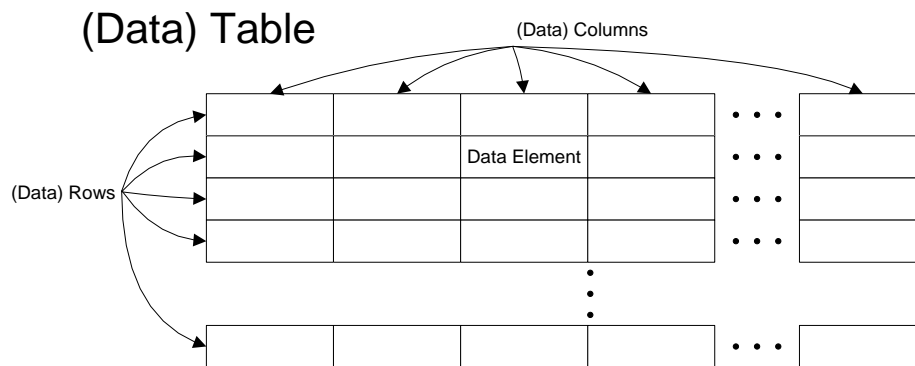
Rule Set – the collection of Rules that currently exists for the Template Output Fields of a particular Period-Type, and thus for all Output Fields for all Input Files for Submittals of that Period-Type.

SBDD ArcGIS File Geodatabase – referred to, for short, as the SBDD Geodatabase, this is the format in which NTIA reporting requirements for a Reporting Period are defined.

Submittal – the Input File(s), Processes, Rules and other data and metadata that pertains to a specific Submittal Type for a specific Provider for a given Reporting Period.

Submittal Type – an NTIA defined database schema for submitting Provider broadband data to the federal government. As of Spring 2012, there are eight defined ESRI feature classes (“Service Address”, “Census Blocks”, “Road Segment”, “Last Mile”, “Middle Mile”, “Wireless Services”, “Service Overview” and “Community Anchor Institutions”).

Table (or Data Table) – a way of regularly arranging a collection of data elements, where each type of data (attribute), irrespective of instance, is a Column and each instance of all Columns is a Row. Spreadsheets provide a familiar way of creating a Table (but only if all cells in a column are of the same type). Databases are composed of groups of Tables. The number of Columns in the Table is fixed by the specification of the attributes that are included. The number of Rows in the Table varies depending on the number of instances (Records or observations) of data that the Table contains. An Input File must always be equivalent to a Data Table no matter what its file format may be.



Template Output Field – a specific attribute, either reported or added, for a Period-Type, which is the basis for creation of an Output Field for an Input File of a Submittal. Since all Submittals for a particular Period-Type have identical reporting requirements, it makes sense to specify these requirements once for each Period-Type. Then, when a particular Submittal is created (or, more specifically, the Input File for a Submittal is created), then the specific Output Fields for this Input File are created from the Template Output Fields. This allows for efficiency in data storage and processing, particularly with regard to Rules.

Template Processing Set – a Processing Set that is saved to act as a Template for a given Period-Type (Reporting Period / Submittal Type). A Template Processing Set is applied (if the check box on the Submittal's tab is checked) when the Processes tab is first visited after loading an Input File's data. If any Processes have been created for an Input File, then the Template Processing Set will be ignored. There can only be one Template Processing Set for each Period-Type. It is created from the Processes tab by saving the current (displayed) Processes and Process Fields (the Processing Set) as a Template; if a Template already existed, it will be replaced. Application of a Template Processing Set will result in a Processing Set that is exactly like the one that was saved (as the basis for the Template) except all Input Fields will be uncoupled (disconnected) and appropriate Input Process Fields will be substituted in their place. All Output Fields, however, will still be coupled. Any Provider Value Processes will be automatically updated to use the attribute values for the active Provider. Thus, the user simply needs to connect the Input Process Fields to this Input File's Input Fields, click "Apply" and "Save", and they're ready to proceed.

Appendix A: User Needs Analysis

DIVA development began with a list of user needs from which the functionality of DIVA was designed and implemented. It should be noted that this project was scoped in July 2010 based on the type and amount of Provider data the development team had seen to that point. As with any software development project, needs have changed somewhat over time. We have done our best to add features to DIVA that respond to insights we gained in subsequent submittals. Following is the original list of requirements.

Functional Area	Req. Number	Requirement
01 - General	1.10	The application should be designed and developed to install and run in Windows XP or Windows 7 (32 or 64 bit).
01 - General	1.20	The application should minimize or eliminate the need for third-party proprietary software that needs to be present on client workstations.
01 - General	1.30	The application should be updateable and configurable by non-programmers.
01 - General	1.40	The application should provide a means to store and recall configuration options for re-use with future input files.
01-General	1.50	The application should provide a means to configure field-specific rules, transformations or other processes that must be performed on each field in the Provider submittal to standardize the output to NTIA specifications.
01 - General	1.60	The application should record and report the results of major data translation and verification/validation operations in a metadata file that would be rendered as a report to the user.
01 - General	1.70	The application should allow the user to select from a set of pre-defined report templates and output formats that define the output of reports.
02 - Data Intake	2.10	The application should be capable of processing all NTIA-acceptable Provider submittal types, including address, road segments, census blocks, service areas and middle/last mile point datasets.
02 - Data Intake	2.20	The application should be able to read and process Excel, delimited text, shapefiles and KML format submittal files from Providers.

Functional Area	Req. Number	Requirement
02 - Data Intake	2.30	The application will be written to assume that input files contain data in a flat-file table schema regardless of the submittal's file type.
02 - Data Intake	2.40	The application should allow users to assign field names to the columns present in Provider submittals that do not contain header rows.
02 - Data Intake	2.50	The application should allow users to "map" the columns found in the Provider files to equivalent SBDD target fields in the SBDD geodatabase schema.
02 - Data Intake	2.60	The application should allow users to identify and document any non-SBDD standard fields for retention in the application's internal database.
02 - Data Intake	2.70	The application should [allow user to capture and] report metadata information which summarizes major items discovered during the initial intake of Provider files.
03- Data Validation	3.10	The application should be capable of "decomposing" a data column value into other data column values (e.g., parse an address into address elements) or "composing" a data column value from other data columns (e.g., assemble a geocodable address).
03- Data Validation	3.20	The application should check for consistency of the Provider Name, DBA Name and FRN in each record.
03- Data Validation	3.30	The application should check the validity and consistency of address information
03- Data Validation	3.40	The application should check the validity and consistency of City and Zipcode information
03- Data Validation	3.50	The application should check the validity and consistency of Census Block identifiers for both Census 2000 and Census 2010. blocks, including the proper area (<= 2 square miles; > 2 square miles)
03- Data Validation	3.60	The application should determine the format that the Min and Max address information has been submitted and standardize the information to report out as AddMin and AddMax fields.

Functional Area	Req. Number	Requirement
03- Data Validation	3.70	The application should check for consistency in how the Speed, Technology and Spectrum Values are coded in a Provider submission.
03- Data Validation	3.80	The application should check for unique records across LocationID, EndUserCat, Spectrum and/or TechID fields
04 - Data Output	4.10	The application should be configurable to adjust to changes in the target NTIA geodatabase.
04 - Data Output	4.20	If the application encounters multiple records for the same technology type and user ID, the application should flag the records that represent the highest maximum download speed.
04 - Data Output	4.30	After processing the Provider input file, the application should output the processed data in a format that can be readily geocoded or joined to the latest NSGIC geodatabase.

Appendix B: Guidelines for Pre-processing Provider Data

Data submitted from Providers will occasionally be in a form or format that DIVA cannot process. For example, information may be broken up into multiple tabs by zip code. Or attribute information, such as Technology Type may be contained in the title of a table and not in the body of the data; or rows of data may be broken by labels indicating an attribute that applies only to the subsequent records. This type of data submittal must be placed into a consistent row-column format in order for DIVA to process it.

Processing Text Files

Open the text file in a text editor or other software that can read and write an ASCII delimited text file. Examine its structure and content. Ensure a clean row, column format and consistent delimiters on each row. Edit the file, as needed, to put the text in a regular format. Using a text editor that can show “hidden” characters, such as tabs and line feeds, is helpful. Notepad++ is one such editor; it is open source and freely downloadable from the Internet.

A user may also import the text file into Excel, Access or another database tool and clean the data there. The file must be saved (exported) in a format that DIVA can process (text, dbf or Excel)

Processing Excel Files

Open the Excel file and examine its structure and content. All data must be on a single tab, in a clean row and column format. At times, this may mean adding a column(s) and moving data from a table title (e.g. Technology Type Asymmetric DSL) into each record of the table (e.g. add a “TechID” column and calculate it to the SBDD value of 10). Save the Excel file in XLS or XLSX formats. Note: DIVA will do translation of values and you may want to wait to convert a textual description of an attribute (e.g. “Asymmetric DSL”) to a value (“10”) until you process in DIVA, as that conversion step will be preserved as part of the processing record.

Processing DBF Files

Open the DBF file in software that can read a DBF. Newer versions of Excel no longer read and write DBF files, but Access does. Examine its structure and content. All data must be on a single tab, in a clean row and column format.

Processing Shapefiles

DIVA does not read shapefiles, per se, but only the attributes stored in the DBF file. It is very easy to corrupt a shapefile by editing its DBF file separately. Therefore, we recommend that (1) you work on a copy of the DBF file and (2) add and unique identifier to the DBF file for later use in joining back to the original shapefile DBF.

To process the DBF's from shapefiles, do the following:

- First make a copy of the shapefile, then add and calculate a UNQID field = ObjectID field of the shapefile.
- Use the copied shapefile's DBF file as the DIVA Input File for the Submittal.
- On the Processes tab in DIVA, copy the UNQID Input Field to the InpSeqNo Output Field (using the Copy Process).
- After processing and exporting to a text file, use the InpSeqNo as a join field back to the UNQID field.

Processing MapInfo Files

Convert the MapInfo files into shapefiles using a freeware convertor. See the above "Processing Shapefiles" for further processing guidelines.

Processing Access Files

Open the Access file and export to a DBF, text or Excel format. Since Access files have to be structured as a database, there should be no other action required.

Appendix C: Object Data Model

A key architectural decision early in the DIVA development process was to use object-relational mapping (ORM) to persist DIVA programming objects. Microsoft's Entity Framework v4.1 (EF) was chosen based primarily on earlier decisions to do code-development using the .NET C# language and the Windows Presentation Foundation (WPF) for the GUI creation. Our "prototype-driven" development cycle made it easy to settle on the "Model-First" approach to using EF. The DIVA Object Data Model (ODM) is simply our class diagram done in EF format.

The version of the ODM for DIVA v1.0 is shown in Figure C-1 below. The following is offered as a short description in how to "read" the diagram. Refer also to the simpler version of this diagram, Figure 1, as you read this explanation.

DIVA entities (objects) are shown as "boxes". The name of the entity (e.g., Provider) is shown at the top of the box. Each entity (e.g., Provider) implies a collection of all such entities (e.g., Providers); the collection name is the same as the data table storing ("persisting") the attribute values with one row in the data table corresponding to one entity. The attributes, or properties, of each object are listed below the name. Note that there will always be at least one attribute, the ID attribute, for each entity. Below the properties are listed "navigational properties"; these are essentially the names of the other entities that are directly related to this object.

Relationships between entities are shown as dashed lines connecting the entities. The "multiplicity" of a relationship is shown by the number next to each end of the relationship "lines"; it is also implied by name assigned to the navigational property (e.g., InputFile has a navigational property named Submittal implying that it can only be related to one Submittal, while the Submittal has a navigational property named InputFiles implying that it can be related to more than one InputFile). Note that a "*" means "many" which means any number (including zero).

The Provider, Reporting Period and Submittal Type entities are "primary" in the sense that they do not depend on other entities (they are the "independent" dimensions of the Submittal "space"). Start with these primary entities when examining the ODM.

The ProviderPeriod entity is an association object linking Provider and Reporting Period. It exists to "host" the Notes attribute. The "multiplicity" shows that there can only be one ProviderPeriod entity at most for every Provider / Reporting Period combination (however, we do not have to have one if not needed to store the

Notes attribute). The “*” on the line next to the Provider means, for example, that a Provider might have many ProviderPeriods associated with it, but there can be only one for each Period (as shown by the other relationship line).

Similarly, the PeriodType entity is an association object linking ReportingPeriod and SubmittalType. It is needed to allow the definition of a SubmittalType (in terms of TemplateOutputFields, etc.) from one Reporting Period to another. The “multiplicity” shows that there can only be one PeriodType entity for every Reporting Period / Submittal Type combination (again, we do not have to have one if that Period / Type combination is not allowed).

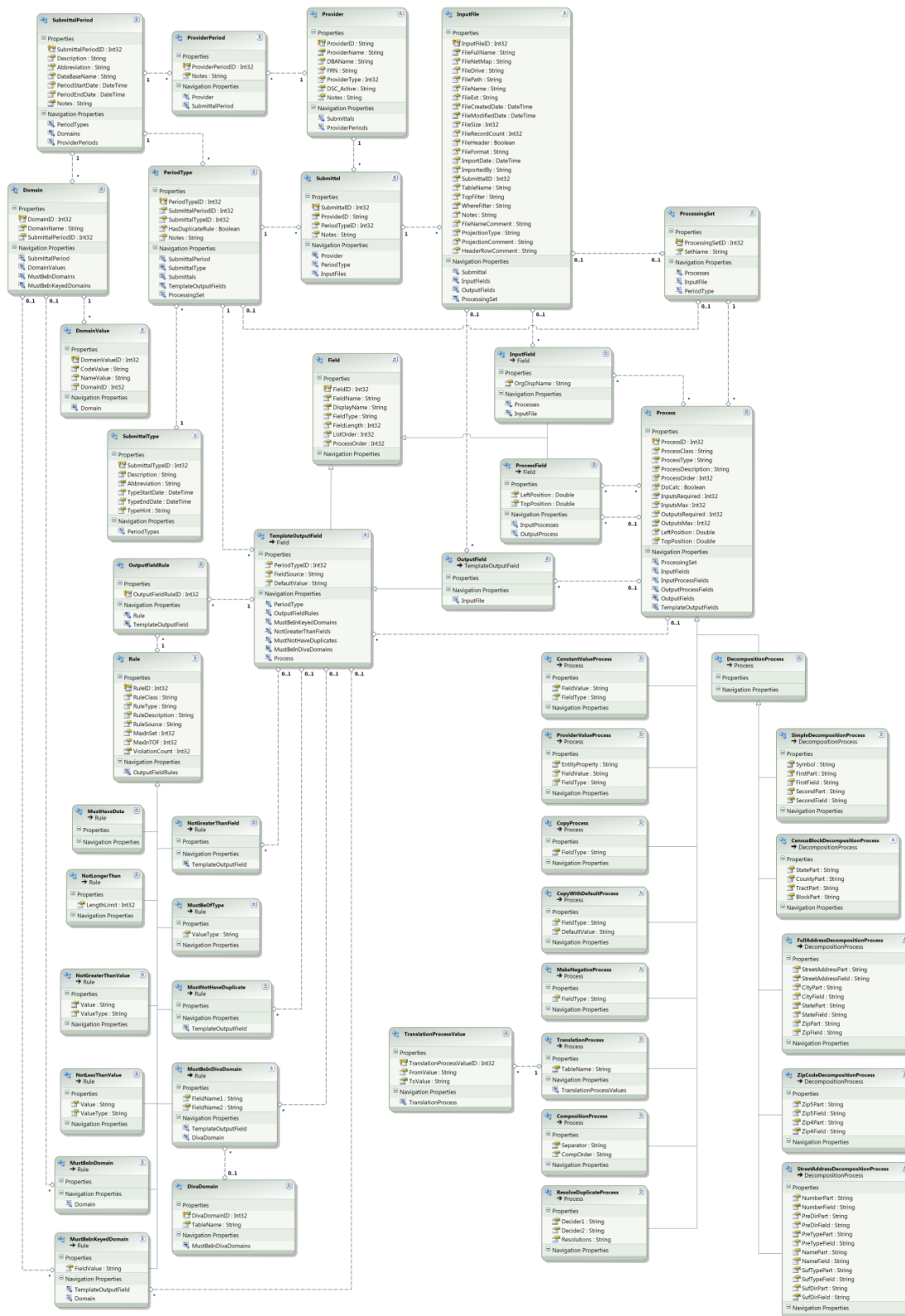


Figure C-1. DIVA v1.0 Object Data Model (ODM)

The TemplateOutputFields are related to the Period-Types; each may belong to only one Period-Type. As explained in the User Manual, these are created, along with the associated Period-Type, during the SBDD XML Importoperation.

There is a many-to-many relationship between Rule and TemplateOutputField entities. This is handled explicitly by creating an association entity called OutputFieldRule between them. There can only be one such OutputFieldRule for every Rule TemplateOutputField combination. This does not, however, prevent us from assigning more than one Rule of a particular type to a TemplateOutputField! We simply have to create a new OutputFieldRule to carry that association. Rules and OutputFieldRules are created, again, by the SBDD XML Importoperation; the user may add additional Rules to TemplateOutputFields on the DIVA Rules tab (an OutputFieldRule entity is automatically created by DIVA each time this is done).

The Submittal entity may only have one Provider and one PeriodType associated with it. The Submittal properties show another possibility. While we have the Provider as a navigational property, we also have the related ProviderID as an attribute of the Submittal. This is done merely for convenience in how we write code to “grab” related entities; we could simply use the navigational property if we wished.

The InputFile entity may only belong to one Submittal (the one on the relationship line next to the Submittal), but each Submittal may have many InputFiles (the “*” on the relationship line next to the InputFile). The ODM makes clear the large number of attributes that DIVA tracks regarding a Provider’s data set.

An InputFile has a ProcessingSet and a collection of InputFields and OutputFields associated with it. These are all created during data loading. Initially the ProcessingSet will be empty, but we will add to it Processes and their associated ProcessFields (if any) as they are created. The InputFields are DIVA’s way of getting a hold of the Provider’s actual data (which sits in another data table, not shown in the ODM). The OutputFields derive from the TemplateOutputFields. Because the Submittal is related to only one Period-Type, it has only those TemplateOutputFields to use in creating its OutputFields, and one OutputField is created for each TemplateOutputField.

In examining the ODM, you will see that there are no “multiplicity” numbers next to some of the lines coming into the TemplateOutputField, InputField and OutputField entities. Instead, there is a line that has a “triangle” on one end of it. These all lead, ultimately, to a Field entity. This is how EF handles “subtyping” objects. The Field entity is actually an “abstract” object; that is, we cannot create a Field object per se. An InputField is a *type-of* Field. A TemplateOutputField is a *type-of* Field. An OutputField, however, is a *type-of* TemplateOutputField. What this means is that an InputField, for example, not only has the properties and navigational properties shown in its box, it also has the properties and navigational properties shown in the box of which it is a type (the Field entity’s box). This allows for efficient storage of commonly shared attributes in the data tables that are used to save (“persist”) these entity attribute values.

A Process is related to a ProcessingSet. It is also, possibly, related to an InputField and / or an OutputField. The multiplicity values make clear that a Process may be related (connected to) more than one InputField, but only one OutputField. A Process may also be related to a ProcessField and this is shown with two relationship lines. One of these lines represents the InputProcessFields navigational property, the other the OutputProcessFields navigational property. It is this set of relationships that allows us to model the connections shown on the Processes tab that are created as we “link” Processes with InputFields and OutputFields, or other Processes through ProcessFields.

A Process is also an “abstract” object; we do not create a Process per se, but a particular *type-of* Process. The various Process entities are shown on the lines that descend from the Process entity. Note that the DecompositionProcess itself is an abstract Process and that we can only make the particular *type-of* DecompositionProcesses that are shown.

Similarly, a Rule is an “abstract” object; we can only make the *type-of* Rules that are shown. So the line leading out of Rules has a triangle on its end.

This takes us through the bulk of the diagram. A few items remain.

Note that the ProcessingSet has two lines leading out of it, each with “0..1” on both ends. This means zero or one related entities. One line leads to the InputFile entity; if this relation is “1” then the ProcessingSet belongs to the InputFile (it is a regular ProcessingSet) and the other relation line will be a “0”. The other line leads to the PeriodType entity; if this relation is “1” then the ProcessingSet belongs to the Period-Type (it is a TemplateProcessingSet) and the other relation line will be a “0”. That is, it has to be one or the other; it cannot be both. And there is no fundamental difference between these two. Note that the Process entity has similar lines leading to OutputField and TemplateOutputField, respectively. Again, a regular ProcessingSet will only have relations going to OutputFields, while a TemplateProcessingSet will only have relations going to TemplateOutputFields.

The Domain entity is attached to the ReportingPeriod entity. Thus, these are specific to a Reporting Period and the same Domains will be used for all Period-Types that are related to that Period (and, thus, all Submittals, Rules through TemplateOutputFields, etc.). The DomainValue entity is related to the Domain that it belongs to; each Domain may have any number of values but a DomainValue may only belong to one Domain.

The TranslationProcessValue is attached to the TranslationProcess. Each TranslationProcess may have many TranslationProcessValues, but each TranslationProcessValue may only belong to one TranslationProcess. This is where the From / To value pairs are stored for this Process.

Finally, the DivaDomain entity is attached to the MustBelInDivaDomain *type-of* Rule entity. The DivaDomain entity does not store its values in another entity (like the Domain does). Instead, it points to a data table that is stored in the DIVA database. The DivaDomain is not specific to a particular Reporting Period and can be used for all Periods if desirable.

Appendix D: Changes DIVA Makes in SQL Server

When DIVA is initially installed, a database named “DIVA” will be created. This database is used to “persist” (save) the state of all DIVA entities (objects). It is managed by Entity Frameworks v4.1 and not by explicit DIVA code. Because of this, you should think carefully before attempting to make any manual changes to DIVA’s data tables and never do so without making a backup beforehand.

When an SBDD geodatabase XML file is imported, a new database will be created specific to that Reporting Period. These Period-specific names follow the pattern “Diva_PData_X##” where “X” is either “F” or “S” (Fall or Spring, respectively) and “##” is the two digit year.

When data is loaded into DIVA’s internal storage, it will be placed in the database appropriate to the Reporting Period and will be given a name that matches the Processing Set’s name. This name has the format “PP_RRR_SS_#” where “PP” is the two character ProviderID attribute value, “RRR” is the Reporting Period’s Abbreviation attribute value, “SS” is the Submittal Type’s Abbreviation attribute value and “#” is the integer value of the Input File’s internal identifier. This name is guaranteed to always be unique.

The DIVA database can be backed up at any time. Usually, this would be done at the end of a Reporting Period before the start of the next Period’s data processing. Restoring the database would thus return DIVA to the state it was in “between” these periods. Because the DIVA database is “uncoupled” from the Period-specific databases, no adjustment would be needed if restoration is done from such backups. Just delete the Period-specific database and re-import the SBDD geodatabase XML file.

Backing up and restoring DIVA during a Period is more complicated in that you need to ensure that only those data tables that should be in the Period-specific database are actually in it. This can be done by backing up the current Period’s database at the same time that DIVA is backed up, and then restore it in synchronization with DIVA restorations.

The Period-specific databases themselves can be backed-up and then removed from SQL Server after a Period is closed and you have no further need of them (e.g., for reporting). They can be restored, if needed, for access to previous Periods.

The DIVA database grows in a fairly “linear” fashion (for each Reporting Period and Submittal added) and should remain relatively small over time. However, the Period-specific databases could be quite large depending on the size and volume of Provider-submitted data sets. In our experience, they tend to be quite a

bit larger than the amount of data added to DIVA during the Reporting Period. The fact that they are “uncoupled” from DIVA greatly eases the “data management” problem.

Note that data tables loaded as DivaDomains are stored in the DIVA database, not a Reporting Period-specific database.